1
Technical Scale of Electrochemistry

Klans Jüttner
Karl-Winnacker Institute, Dechema e.v., Frankfurtan Main, Germany

1.1 **Introduction** **3**

1.2 **Advantages and Disadvantages of Electrochemical Processes** **4**

1.3 **Electrochemical Reaction Engineering** **5**

1.4 **Fundamentals** **6**
1.4.1 Current Efficiency 7
1.4.2 Space-time Yield 8

1.5 **Electrochemical Thermodynamics** **8**
1.5.1 Energy Yield 11
1.5.2 Specific Electrical Energy Consumption 12

1.6 **Electrochemical Cell Design** **12**
References 19

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

1.1 Introduction

From an academic point of view, electrochemistry has become a written chapter in science history, manifested, for example, by the Nernst equation or the Butler–Volmer equation, to mention a few of the most common relations. The interest of the academic community in electrochemistry as a discipline of fundamental importance has gradually declined in recent years. Electrochemistry today presents itself as a strongly diversified field, which one would call ''applied electrochemistry'', ranging from the study of electrode processes for synthesis of inorganic or organic compounds, energy conversion, sensor development, electroanalytic devices, electrolytic corrosion to solid-state ionics, and any kind of system where charge transfer is involved at an electrified interface between electronic and ionic conductors. Research is more and more specialized on very particular problems, often of a multidisciplinary character, being in touch with biology, medicine, microelectronics, structuring, and functionalizing of surfaces at the nanoscale as a part of materials and surface science.

Another view on ''applied electrochemistry'' is that of the electrochemical engineer. Electrochemical engineering has become an established discipline of its own, which deals with the description and optimization of electrochemical processes based on the fundamental laws of chemistry and electrochemistry. Essentially, it is an interdisciplinary field, which came up in the 1960s after the development of chemical engineering, providing the tools of unit operations for the systematic description and treatment of chemical processes in a way that linked the underlying physical and physicochemical principles to chemical technology. One of the pioneers was C. Wagner, who initiated this development by his article ''The Scope of Electrochemical Engineering'' in ''Advances in Electrochemistry and Electrochemical Engineering'' [1]. It soon became apparent that this kind of treatment can be adopted and transferred to the field of electrochemical reaction and process engineering [1–19]. Today, established technical processes include chloralkali electrolysis and related processes of hypochlorite and chlorate formation, electrowinning of metals like aluminum and magnesium from molten salt electrolytes, hydrometallurgy for electrowinning of metals like copper, nickel, zinc, and refining of cast metals, electro-organic synthesis of, for example, adipodinitrile as precursor for polyamide (nylon) production, electroplating in the

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

galvanic industry, or electrophoretic coating in the automotive industry. The environmental aspect has become a major issue and a crucial factor in the development of industrial processes to meet the requirements of sustainable development. In this field, electrochemistry offers promising approaches due to its environmental compatibility and use of the electron as a "clean reagent" [20–28]. There is common agreement among scientists that electrochemically based processes will be of increasing importance in the future to meet the economic and social challenges resulting from urgent demands of low-grade raw materials' utilization, energy savings, and protection of the environment.

1.2 Advantages and Disadvantages of Electrochemical Processes

Applied electrochemistry can provide valuable cost efficient and environmentally friendly contributions to industrial process development with a minimum of waste production and toxic material. Examples are the implementation of electrochemical effluent treatment, for example, the removal of heavy metal ions from solutions, destruction of organic pollutants, or abatement of gases. Further progress has been made in inorganic and organic electrosynthesis, fuel cell technology, primary and secondary batteries, for example, metal-hydride and lithium-ion batteries. Examples of innovative industrial processes are the membrane process in the chloralkali industry and the implementation of the gas-diffusion electrode (GDE) in hydrochloric acid electrolysis with oxygen reduction instead of hydrogen evolution at the cathode [28]. Figure 1 shows the cell room of the electrolyzer at Bayer Materials Science AG in Brunsbüttel, Germany.

The main advantages of electrochemical processes are as follows:

- *Versatility*: Direct or indirect oxidation and reduction, phase separation, concentration or dilution, biocide functionality, applicability to a variety of media and pollutants in gases, liquids, and solids, and treatment of small to large volumes from microliters up to millions of liters.
- *Energy efficiency*: Lower temperature requirements than their nonelectrochemical counterparts, for example, anodic destruction of organic pollutants instead of thermal incineration; power losses caused by inhomogeneous current distribution, voltage drop, and side reactions being minimized by optimization of electrode structure and cell design.
- *Amenability to automation*: The system inherent variables of electrochemical processes, for example, electrode potential and cell current, are particularly suitable for facilitating process automation and control.
- *Cost effectiveness*: Cell constructions and peripheral equipment are generally simple and, if properly designed, also inexpensive.

Since electrochemical reactions take place at the interface of an electronic conductor – *the electrode* – and an ion conductor – *the electrolyte* – electrochemical processes are of heterogeneous nature. This implies that, despite the advantages mentioned earlier, the performance of electrochemical processes suffers from mass-transport limitations and the size of the specific electrode area. A crucial point is also the stability of cell components in



Fig. 1 Cell room of the hydrochloric acid electrolyzer with GDE in Brunsbüttel [28] (photo: Bayer Material Science AG).

contact with aggressive media, in particular, the durability and long-term stability of the electrode material and the catalysts.

1.3 Electrochemical Reaction Engineering

Electrochemical reaction engineering deals with modeling, computation, and prediction of production rates of electrochemical processes under real technical conditions in a way that technical processes can reach their optimum performance at the industrial scale. As in chemical engineering, it centers on the appropriate choice of the electrochemical reactor, its size and geometry, mode of operation, and the operation conditions. This includes calculation of performance parameters, such as space-time yield, selectivity, degree of conversion, and energy efficiency for a given reaction in a special reactor.

Electrochemical thermodynamics is the basis of electrochemical reaction engineering to estimate the driving forces and the energetics of the electrochemical reaction. Predictions of how fast the reactions may occur are answered by electrochemical kinetics. Such predictions need input from fundamental investigations of the reaction mechanism and the microkinetics at the molecular scale [29–50]. While microkinetics describes the charge transfer under idealized conditions at the laboratory scale, macrokinetics deals with the kinetics of the large-scale process for a certain cell design, taking into account the various transport phenomena. The microkinetics combined with calculations of the spatial distribution of the potential

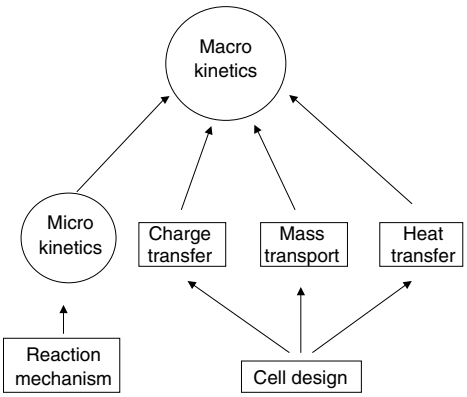


Fig. 2 Concept of microkinetics and macrokinetics for electrochemical reactions.

and current density at the electrodes leads to the macrokinetic description of an electrochemical reactor [3–5, 7, 9, 10, 12–18]. The relation between micro- and macrokinetics is shown schematically in Fig. 2.

The phenomenon of charge transport, which is unique to all electrochemical processes, must be considered along with mass, heat, and momentum transport. The charge transport determines the current distribution in an electrochemical cell, and has far-reaching implications on the current efficiency, space-time yield, specific energy consumption, and the scale-up of electrochemical reactors.

1.4 Fundamentals

In the following, similarities and differences in the fundamental description of chemical and electrochemical reactions are considered with a definition of symbols and signs [51, 52].

In general, a chemical reaction is written as the sum of educts A_i and products B_j

$$\sum_i \nu_{A_i} A_i = \sum_j \nu_{B_j} B_j \qquad (1)$$

with stoichiometric coefficients ν_i. The rate of the reaction is defined by

$$r = \frac{1}{\nu_{A_i}} \frac{dn_{A_i}}{dt} = \frac{1}{\nu_{B_j}} \frac{dn_{B_j}}{dt} \quad \forall \quad i, j. \quad (2)$$

where the stoichiometric coefficients are by definition negative for the educts, $\nu_{A_i} < 0$, and positive for the products, $\nu_{B_j} > 0$. Taking into account that the mole numbers of the educts are decreasing ($dn_{A_i}/dt < 0$) and those of the products are increasing ($dn_{B_i}/dt > 0$), the reaction rate r (mol sec^{-1}) is always a positive number ($r > 0$).

Electrochemical reactions are characterized by at least one electron charge transfer step taking place at the electrode/electrolyte interface. Therefore, the electron e^- appears as a reactant in the reaction equation with its stoichiometric coefficient ν_e

$$\sum_i \nu_{A_i} A_i + \nu_e e^- = \sum_j \nu_{B_j} B_j, \qquad (3)$$

Typical electrode reactions are given in Table 1.

Applying Eq. (2) to the reactant e^-, the reaction rate normalized to the geometrical

Tab. 1 Electrode reactions and standard potentials E^o [18]

Electrode reaction	*Standard potential E^o/V vs SHE*
$Li^+ + e^- \rightarrow Li$	−3.045
$Na^+ + e^- \rightarrow Na$	−2.711
$Al^{3+} + 3e^- \rightarrow Al$	−1.706
$2H_2O + 2e^- \rightarrow H_2 + 2OH^-$	−0.8277
$Zn^{2+} + 2e^- \rightarrow Zn$	−0.763
$Fe^{2+} + 2e^- \rightarrow Fe$	−0.409
$Ni^{2+} + 2e^- \rightarrow Ni$	−0.230
$Pb^{2+} + 2e^- \rightarrow Pb$	−0.126
$2H_3O^+ + 2e^- \rightarrow H_2 + 2H_2O$	0
$Cu^{2+} + 2e^- \rightarrow Cu$	0.340
$[Fe(CN)_6]^{3-} + e^- \rightarrow [Fe(CN)_6]^{4-}$	0.356
$O_2 + 2H_3O^+ + 2e^- \rightarrow H_2O_2 + 2H_2O$	0.682
$Fe^{3+} + e^- \rightarrow Fe^{2+}$	0.771
$Ag^+ + e^- \rightarrow Ag$	0.799
$O_2 + 4H_3O^+ + 4e^- \rightarrow 6H_2O$	1.229
$Cl_2 + 2e^- \rightarrow 2Cl^-$	1.37
$Ce^{4+} + e^- \rightarrow Ce^{3+}$	1.443
$MnO_4^- + 8H_3O^+ + 5e^- \rightarrow Mn^{2+} + 12H_2O$	1.491
$O_3 + 2H_3O^+ + 2e^- \rightarrow O_2 + 3H_2O$	2.07
$F_2 + 2e^- \rightarrow 2F^-$	2.87

electrode surface area A_e can be written as

$$r = \frac{1}{A_e \nu_e} \frac{dn_e}{dt} \tag{4}$$

The number of moles n_e transferred within a certain time t can be expressed by the electrical charge $Q_t = It = A_e it$ and the Faraday constant $F = 96\,485\ \mathrm{As\,mol^{-1}}$

$$n_e = \frac{Q_t}{F} \tag{5}$$

The linear relation

$$r = \frac{i}{\nu_e F}, \tag{6}$$

which results from (4) and (5) shows that the rate of an electrochemical reaction r ($\mathrm{mol\,cm^{-2}\,sec^{-1}}$) is proportional to the current density i ($\mathrm{A\,cm^{-2}}$), the sign of which is determined by the sign of the stoichiometric coefficient ν_e. It follows that the current density is negative (cathodic), if the electrode reaction describes a reduction process (e.g., Eq. 3), and positive (anodic) for an oxidation process (opposite direction).

Equation (6) is the differential form of Faraday's law. The common expression is obtained by integration of Eq. (2)

$$\frac{1}{\nu_P} \frac{dn_P}{dt} = \frac{1}{\nu_e} \frac{dn_e}{dt} \tag{7}$$

when applied to the simplified electrode reaction $\nu_R R + \nu_e e^- = \nu_P P$ with a single reactant R and product P

$$m_P = \frac{\nu_P M_P}{\nu_e F} Q_t \quad \text{or with respect to the reactant R} \quad m_R = \frac{\nu_R M_R}{\nu_e F} Q_t \tag{8}$$

Here m_P and m_R are the mass and M_P and M_R the molar mass of R and P, respectively. Owing to the sign convention of ν_i, the mass m_R in Eq. (8) is negative, indicating that R is consumed during the reaction ($Q_t < 0$).

A prerequisite for the validity of these equations is that only the considered reaction takes place at the electrode. In reality, side reactions can occur in parallel, the individual contribution of which is given by their current efficiency Φ^e.

1.4.1 Current Efficiency

The classical definition of the current efficiency Φ^e is based on Faraday's law and relates the mole numbers n_P of the product to that of the electron n_e consumed in the

electrochemical reaction.

$$\Phi_{\mathrm{P}}^{\mathrm{e}} = \frac{m_{\mathrm{P}} \nu_{\mathrm{e}} F}{\nu_{\mathrm{P}} M_{\mathrm{P}} Q_{\mathrm{t}}} = \frac{n_{\mathrm{P}}/\nu_{\mathrm{P}}}{n_{\mathrm{e}}/\nu_{\mathrm{e}}} \qquad (9)$$

In that sense, Φ^{e} is the "yield of charge" Q_{t} consumed in the course of the reaction. The current efficiency defined by Eq. (9) represents an average value. The differential or point current efficiency is defined by [17]

$$\Phi_{\mathrm{P}}^{\mathrm{e}'} = \frac{1}{\nu_{\mathrm{P}}} \frac{\mathrm{d}n_{\mathrm{P}}}{\mathrm{d}t} \Big/ \frac{A_{\mathrm{e}} i}{\nu_{\mathrm{e}} F} \qquad (10)$$

which is consistent with the definition of the partial current density of the product P, $i_{\mathrm{P}} = i\,\Phi_{\mathrm{P}}^{\mathrm{e}'}$.

1.4.2 Space-time Yield

Another performance criterion, which is of particular importance for the calculation of capital investment and running costs of an electrochemical reactor is the space-time ρ_{st} yield, which is the amount of product m_{P} produced per unit of time t and reactor volume V_{r}

$$\rho_{\mathrm{st}} = \frac{1}{V_{\mathrm{r}}} \frac{\mathrm{d}m_{\mathrm{P}}}{\mathrm{d}t} \qquad (11)$$

Substitution of m_{P} from Eq. (8) leads to the point space-time yield

$$\rho_{\mathrm{st}} = \frac{\nu_{\mathrm{P}} M_{\mathrm{P}}}{\nu_{\mathrm{e}} F} \Phi^{\mathrm{e}} a_{\mathrm{e}} i_{(\mathrm{t})} \qquad (12)$$

which contains a design factor (the specific electrode area $a_{\mathrm{e}} = A_{\mathrm{e}}/V_{\mathrm{r}}(\mathrm{cm}^2/\mathrm{cm}^3)$) and a reaction factor (the product of current density and current efficiency). It is evident that optimization will be directed toward enlargement of a_{e} and increase in the reaction rate, that is, the current density i. The latter can be achieved, for example, by shifting the reaction from charge transfer to a mass-transport-controlled regime, where the diffusion-limited current density, as the maximum of the reaction rate, can be reached

$$i_{\mathrm{d,lim}} = -\frac{\nu_{\mathrm{e}}}{\nu_{\mathrm{R}}} F\, k_{\mathrm{m}} c_{\mathrm{R}} \qquad (13)$$

This leads to the following expression:

$$\rho_{\mathrm{st}} = -\frac{\nu_{\mathrm{P}}}{\nu_{\mathrm{R}}} M_{\mathrm{P}} \Phi^{\mathrm{e}} a_{\mathrm{e}} k_{\mathrm{m}} c_{\mathrm{R}} \qquad (14)$$

In this case, ρ_{st} is proportional to the mass-transport coefficient k_{m}, which can be increased by forced convection, for example, through agitation of the solution, enhanced electrolyte flow, turbulence promoters, or setting the electrodes in motion. Equations (12) and (14) are the key formula for the optimization of electrochemical cell design. Equation (14) is of particular importance for reactions with low concentrations, for example, recycling of metals from spent solutions or wastewater.

1.5 Electrochemical Thermodynamics

The Galvani potential difference arising at the electrode/solution interface, $\Delta\varphi = \varphi_{\mathrm{s}} - \varphi_{\mathrm{e}}$, depends specifically on the energy and density of electronic states in the two phases in contact. In principle, it is impossible to measure $\Delta\varphi$ at a single electrode. By combining two electrodes 1 and 2 to an electrochemical cell, the difference between the two interface potentials can be measured as the cell voltage $U_{\mathrm{c}} = \Delta\varphi_1 - \Delta\varphi_2$. To normalize this potential measurement, the standard hydrogen electrode (SHE) is used as a reference and the cell voltage U_{c} is denoted as electrode potential E versus SHE. The

electrode potential at equilibrium, E_o, depends on the concentrations c_i or activities a_i of the reactants according to the Nernst equation

$$E_o = E^o + \frac{RT}{\nu_e F} \sum_i \nu_i \ln a_i \qquad (15)$$

where E^o is the standard potential at $a_i = 1$, $T = 298$ K, and pressure $p = 1.013 \times 10^5$ Pa $= 1013$ bar. Table 1 contains the standard potentials E^o of the selected electrode reactions. The standard potentials E^o versus SHE or cell voltages U_c^o can be calculated from thermodynamic data.

The standard Gibbs free energy of the overall cell reaction ΔG_{298}^o is related to the electrical work nFU_c^o,

$$\Delta G_{298}^o = -nF\ U_c^o \qquad (16)$$

which can be gained under quasireversible operating conditions of the cell, that is, at an infinitely low reaction rate.

ΔG_{298}^o is the stoichiometric sum of the Gibbs free energies of formation, $\Delta G_{f,i}^o$, of the reactants involved in the reaction

$$\Delta G_{298}^o = \sum \nu_i \Delta G_{f,i}^o \qquad (17)$$

In (16), n is the number of electrons transferred in the overall process to maintain electroneutrality. Thermodynamic data for many chemical reactants and compounds are available as standard enthalpies $\Delta H_{f,i}^o$ and entropies $S_{f,i}^o$ from thermodynamic tables in handbooks, for example, [53, 54]. For the elements and for the proton H^+ (aq) in aqueous solution (H_3O^+), the $\Delta H_{f,i}^o$ values are zero by definition. Thermodynamic data for some typical electrochemical reactants are given in Table 2.

These data are related to the Gibbs free energy of reaction ΔG_{298}^o by the Gibbs–Helmholtz equation

$$\Delta G_{298}^o = \Delta H_{298}^o - T\Delta S_{298}^o \qquad (18)$$

where $T\Delta S^o$ is the reversible heat exchanged with the environment. The entropy change ΔS^o can be either positive or negative and is directly related to the temperature coefficient of the cell

$$\left(\frac{\partial U_c^o}{\partial T}\right)_p = -\frac{1}{nF}\frac{\partial \Delta G_{298}^o}{\partial T} = \frac{1}{nF}\Delta S_{298}^o \qquad (19)$$

Tab. 2 Thermodynamic data of typical reactants [41]

Reactant	*State of aggregation*	ΔH_{298}^o [kJmol^{-1}]	S_{298}^o [JK^{-1}mol^{-1}]
H_2	g	0	130.74
Cl_2	g	0	223.09
H_3O^+	aq	0	0
Cl^-	aq	−167.54	55.13
O_2	g	0	205.25
H_2O	l	−285.25	70.12
Zn^0	s	0	41.65
Zn^{++}	aq	−152.51	−106.54
HCl	g	−92.53	186.79
C	s	0	5.69
CO	g	−110.5	198.0

(g: gas; l: liquid; s: solid; aq: aqueous solution of activity one)

The ΔG^{o}_{298} values are calculated from the stoichiometric sum of the enthalpies of formation $\Delta H^{o}_{f,i}$ and the entropies $S^{o}_{f,i}$ of the species involved

$$\Delta G^{o}_{298} = \sum_i \nu_i \Delta H^{o}_{f,i} - T \sum_i \nu_i S^{o}_{f,i} \tag{20}$$

If ΔG^{o}_{298} is negative, the reaction is exergonic and should occur spontaneously, otherwise it is endergonic and does not occur or occurs in the reverse direction only. The negative sign in Eq. (16) implies that the cell voltage or standard potential of an exergonic reaction must be positive. This calls for a definition of the cell voltage measurement. According to IUPAC [51], the cell voltage is the potential difference between the right half cell I and the left half cell II, as shown in Fig. 3. The direction of the electronic current in the outer electronic circuit is defined from left to right, which means that cathodic reduction must take place in half cell I and anodic oxidation in half cell II.

This is demonstrated by combining an oxygen electrode and a hydrogen electrode as the two half cells of an electrochemical cell. The half cell reactions are commonly formulated as reductions:

Oxygen electrode:

$$\frac{1}{2}O_2 + 2H^+ + 2e^- = H_2O \quad \text{with}$$
$$E^{o}_{O_2/H_2O} = +1.229\ \text{V versus SHE} \tag{21}$$

Hydrogen electrode:

$$H^+ + e^- = \frac{1}{2}H_2 \quad \text{with} \quad E^{o}_{H^+/H_2} = 0.0\ \text{V versus SHE} \tag{22}$$

If the oxygen electrode is placed on the right and the hydrogen electrode on the left in Fig. 3, according to IUPAC convention, the electrode reactions must be written

$$\frac{1}{2}O_2 + 2H^+ + 2e^- = H_2O \tag{23}$$

$$H_2 = 2H^+ + 2e^- \tag{24}$$

so that the sum yields the overall cell reaction and the electrons are cancelled.

$$\frac{1}{2}O_2 + H_2 = H_2O \tag{25}$$

For the example considered, the overall electron-transfer number in Eq. (16) is $n = 2$.

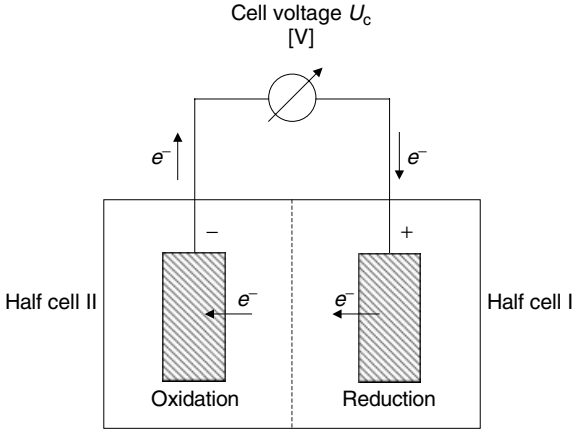


Fig. 3 Electrochemical cell for the definition of cell voltage measurement according to the IUPAC convention [51].

The cell voltage under standard conditions is the potential difference between the right and left half cell: $U_c^o = E^o_{O_2/H_2O} - E^o_{H^+/H_2} = 1.229$ V. Since U_c^o is positive, the Gibbs free energy of the reaction is negative according to Eq. (16)

$$\begin{aligned}\Delta G^o_{298} &= -2.96485 \\ &\quad \cdot (1.299 - 0)\text{VAs mol}^{-1} \\ &= -237.160\ \text{kJ mol}^{-1}.\end{aligned}$$

The cell reaction (25) occurs spontaneously and the cell delivers electrical work nFU_c^o, which is utilized in H_2/O_2 fuel cells (see also Chapter 8). Calculation of ΔG^o_{298} from thermodynamic data in Table 2 yields

$$\begin{aligned}\Delta G^o_{298} &= (\Delta H^o_{H_2O_{fl}} - \Delta H^o_{H_{2,g}} - \frac{1}{2}\Delta H^o_{O_{2,g}}) \\ &\quad - T(S^o_{H_2O_{fl}} - S^o_{H_{2,g}} - \frac{1}{2}S^o_{O_{2,g}})\end{aligned} \tag{26}$$

$$\begin{aligned}\Delta G^o_{298} &= (-285.25 - 0 - 0) - 298 \\ &\quad \cdot (70.12 - 130.74 - \frac{1}{2}205.25) \\ &= -236.603\ \text{kJ mol}^{-1}\end{aligned}$$

The reversible heat $T\Delta S^o = -48.647$ kJ mol^{-1} is negative. It is released by the system and cannot be exploited as electrical work.

If the two half cells are arranged opposite with the hydrogen electrode on the right and the oxygen electrode on the left, the corresponding reactions are reversed and the overall process refers to the electrolysis of water with hydrogen and oxygen as the products. In this case, the cell voltage U_c^o is negative and ΔG^o_{298} is positive. The reversible heat $T\Delta S^o$ is positive and is taken up from the environment; the cell cools down in the course of the reaction. This effect can be compensated if the cell is operated at an increased cell voltage generating the corresponding amount of Joule's heat in the system. The corresponding cell voltage is denoted as the thermoneutral voltage

$$\begin{aligned}U_{th} &= -\frac{1}{nF}\Delta H^o_{298} \quad \text{with} \\ \Delta H^o_{298} &= \Delta G^o_{298} + T\Delta S^o_{298}.\end{aligned} \tag{27}$$

It is interesting to note that, in contrast to chemical reactions, electrochemical cells can operate in principle against a positive ΔG if an outer current source is applied, driving the electrons in the opposite direction. This is the principle of the secondary batteries and electrochemical power sources, which can be charged and discharged to accumulate and release electrical energy.

1.5.1 Energy Yield

The costs of electrochemical processes are closely connected to the energy efficiency or energy yield, which is either related to the change of the Gibbs free energy of the reaction ΔG

$$\gamma_G = \frac{\Delta G \Phi^e}{nFU_c} = \frac{U_{c,o}}{U_c}\Phi^e, \tag{28}$$

or to the thermal energy (enthalpy change ΔH)

$$\gamma_H = \frac{\Delta H \Phi^e}{nFU_c} = \frac{U_{th}}{U_c}\Phi^e. \tag{29}$$

Here $U_{c,o} = -\Delta G/nF$ is the cell voltage at equilibrium, U_c is the actual cell voltage under current operation, and U_{th} is the thermoneutral voltage, Eq. (27), which permits isothermal operation.

1.5.2 Specific Electrical Energy Consumption

For solid or liquid products, the specific electrical energy consumed is usually mass related

$$E_s^e = \frac{U_c \nu_e F}{\nu_P M_P \Phi^e} (\mathrm{J\ kg^{-1}}) \qquad (30)$$

1.6 Electrochemical Cell Design

Different electrochemical cells have been developed for various applications in the past. According to the criteria of high mass-transport coefficient k_m and/or large specific electrode area a_e, as predicted by Eq. (14), these cells can be classified as [20]:

- Cells with relatively small electrode area but *improved mass transport* to increase k_m by setting the electrodes in motion or by applying turbulence promoters. Examples are the pump cell [55, 56], the Chemelec cell [57], the ECO cell [58–60], the beat-rod cell [61, 62], and cells with vibrating electrodes or electrolytes [63].
- Cells with *enlarged electrode area* in a small cell volume are found in the multiple-cathode cell [64], the Swiss-role cell [65], the extended surface electrolysis (ESE) cell [66], and the capillary gap cell [13, 67].
- Cells with *three-dimensional electrodes* providing *enlarged specific electrode area* and *improved mass transport* due to the specific fluid dynamics inside the three-dimensional structure are, for example, the porous flow-through cell [68], the RETEC (RETEC is a trademark of ELTECH Systems Inc., Cardon, Ohio) cell [15], the packed-bed cell [69–71], the fluidized-bed cell [72–76], and the rolling tube cell [77, 78].

Some of them are shown, for example, in Fig. 4(a–c).

The *pump cell* is the rotating analogue to the capillary gap cell [67]. Although this cell has not yet been employed on an industrial scale, it is interesting because of its operation principles. It consists of two static disc electrodes with current connections and a bipolar rotating disk in between. The electrolyte is pumped radially from the central tube to the outer circumference by the rotating disk. The advantage of this design is that the mass-transfer coefficient can be controlled independent of the flow rate and the residence time of the electrolyte.

The *Chemelec cell*, Fig. 4(b), uses a fluidized bed of glass spheres as turbulence promoters to improve the mass transport to the electrodes consisting of a series of closely spaced gauze or expanded metal sheet. The residence time and the degree of conversion per pass are relatively low because the flow rate has to exceed the fluidization velocity of the bed. This cell is therefore suitable for pretreatment or recycling operations and is commonly used in the electroplating industry for maintaining a moderate metal ion concentration in a recirculated wash-water tank.

High mass-transport coefficients are obtained in cells with a *rotating cylinder electrode* (RCE) and a small gap between the anode and the cathode, Fig. 4(a). High rates of mass transport are experienced in the turbulent flow regime, so that RCE reactors allow metal deposition at high speed, even from dilute solutions. RCE reactors have been operated at a scale involving diameters from 5 to 100 cm, with rotation speeds from 100 to 1500 rpm and currents from 1 A to 10 kA [79]. It

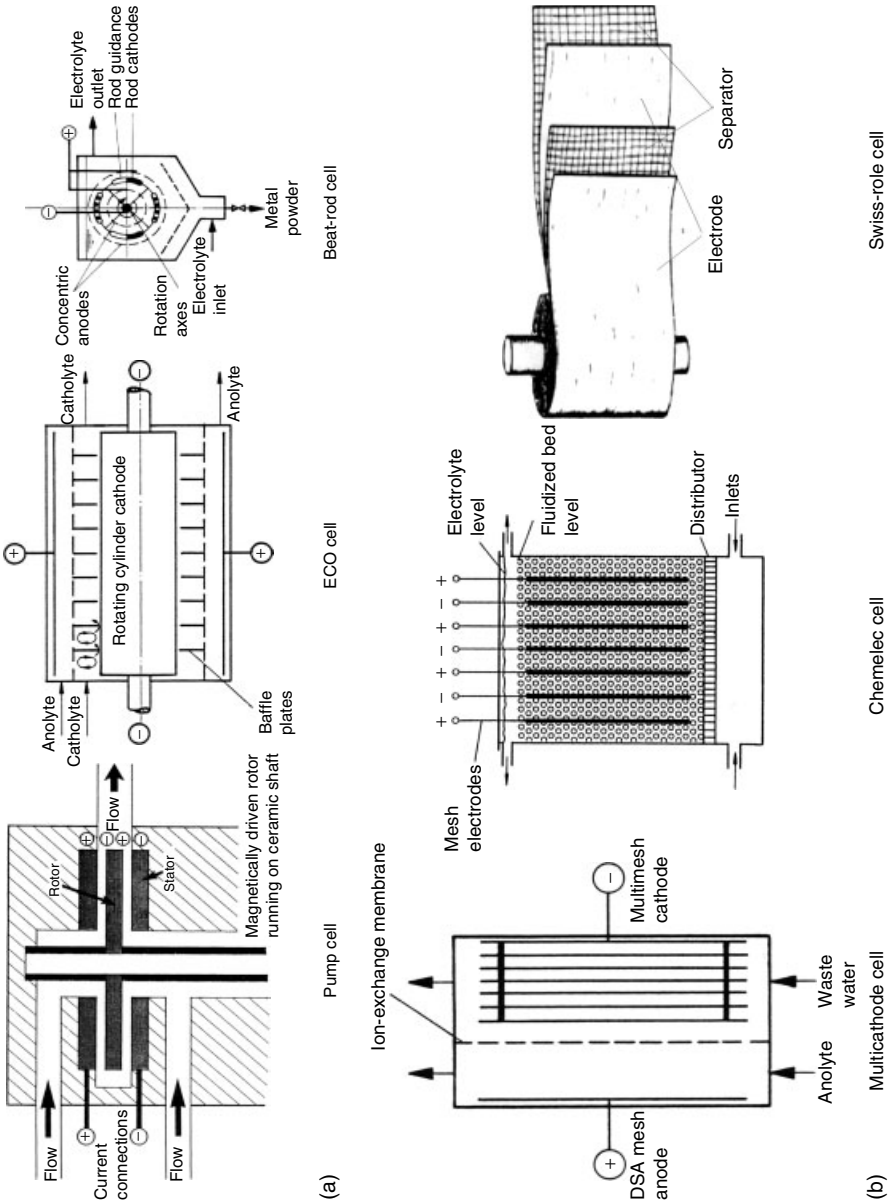
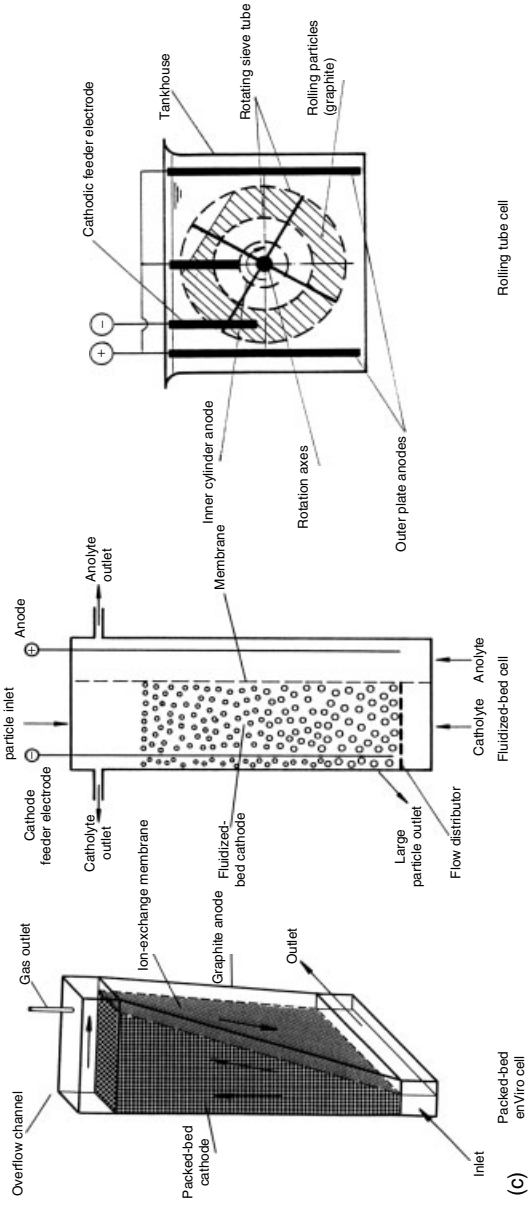


Fig. 4 (a–c) Electrochemical cell design [13].



(c)

Fig. 4 *(Continued).*

is possible to design RCE reactors with a scrapping device to remove deposited metal powder continuously from the cathode surface. The so-called ECO cell, Fig. 4(a), with a cascade of baffle plates around the rotating cylinder cathode, which prevent back mixing of the solution to a great extent, attains high degrees of conversion.

A cell with a special design is the *beat-rod cell*, Fig. 4(a), which was mainly used in small electroplating shops. Supported cathode rods slowly rotate within an annular type of chamber in the electrolysis tank. The rods strike one another so that the metal deposit is removed and settles as a powder at the bottom.

An efficient way of fitting large planar electrodes in a small volume is the design of the *Swiss-role cell*, Fig. 4(b). Thin metal foils are separated by a plastic mesh and wrapped around a central core. The ohmic drop within the thin metal foil is largely compensated by contacting the foils on the opposite side. The electrolyte flows axially through the electrode pack. Large space-time yields are realized and mass-transport conditions are favorable since the separator also functions as a turbulence promoter. Mesh electrodes wrapped around a perforated winding core lead to another version of the Swiss-role cell with radial flow. Cells with mesh electrodes instead of foils are known as extended surface electrode *ESE cells*. A commercial system of this type is the *RETEC cell*, which contains 6–50 three-dimensional cathodes as flow-through metal sponge electrodes [15]. The cell is used in closed water recycling systems in the electroplating industry with applications similar to those of the *Chemelec cell*, Fig. 4(b).

A simple three-dimensional electrode is obtained by using a packed bed of conductive particulate material through which the electrolyte is pumped. Numerous versions of *packed-bed cells* have been described in the literature [13, 68, 80–84]. The potential distribution and optimum bed depth for metal deposition, $\mathrm{M}e^{z+} + ze^- = \mathrm{M}e$, was calculated as a function of the effective particle and solution conductivities (κ_p, κ_s) [13, 20]. For rectangular cell geometry and fully developed diffusion limited current density within the particle bed, the following relation was found for the optimal bed depth:

$$h_\mathrm{opt} = \sqrt{\frac{2\varepsilon\kappa_\mathrm{s}\Delta\eta}{a_\mathrm{e} z F k_\mathrm{m} c_{\mathrm{M}e^{z+}}}} \qquad \text{(condition } \kappa_\mathrm{p} \gg \kappa_\mathrm{s}\text{)} \qquad (31)$$

where ε is the voidage of the 3-D electrode structure and $\Delta\eta$ the overvoltage range of the prevailing limiting diffusion current density in the microkinetic polarization curve. According to Eq. (31), the optimal bed depth increases with decreasing metal ion concentration by an inverse square root dependence. This has led to the *enViro-cell* design depicted in Fig. 4(c), where the bed depth widens with decreasing concentration along the flow direction of the solution.

The principle of *fluidized-bed electrolysis* originates from Fleischmann and Goodridge [73]. It has been examined for different applications [72, 85, 86] and was applied on an industrial scale for metal recovery [75, 76]. The electrolyte flows from bottom to top through a bed of fluidized particles, which are charged via a feeder electrode. The advantage is that the particles are held in suspension, thus avoiding bed blockage by the deposited metal. Because the height of the cell is restricted to about 2 m for hydraulic reasons and the fluidization velocity has to be maintained at a relatively high level, short residence times and thus only a limited degree of

conversion per pass can be achieved. Continuous operation with recirculation of the solution and cascade arrangement of cells has therefore been used in practical applications.

Moving and circulating bed electrodes, also denoted as a spouted bed, are described in the literature [87]. As an example, the *rolling tube cell* is shown in Fig. 4(c) [77, 78]. In principle, this cell resembles the well-known plating process for piece goods, using slowly rotating barrels. The rotating drum is only partially filled with particles to achieve thorough agitation. This cell has proved to be especially useful for silver and gold recovery. Further developments of moving bed electrodes and their application can be found elsewhere [88].

In Fig. 5, the specific electrode area a_e (cm^{-1}) of various electrochemical cells is shown as a function of a characteristic length (interelectrode distance or particle diameter). It can be seen that the electrochemical cells differ significantly in their specific area and characteristic length by orders of magnitude. At a current efficiency of $\Phi^e = 1$, the space-time yield is mainly determined by the product

$$a_e i = \frac{I}{V_r} = j^*. \qquad (32)$$

The quantity j^* ($A\,cm^{-3}$) can be considered as a volumetric current density or a current concentration, also shown in Fig. 5 as a function of the characteristic length. Extremely high current densities, for example, 10 $kA\,cm^{-2}$ in the chloralkali electrolysis, can be applied only in a limited number of cases, the number of cells with high specific electrode area is restricted to cases of low current densities [13].

A special cell concept combining ease of construction and scale-up is the *capillary gap cell*, which is mainly applied in electro-organic synthesis, where low conductivities of the electrolytes are a major problem, Figs 6 and 7, [13, 28, 67, 89]. The cell consists of circular disk electrodes with a small interelectrode gap (1–2 mm) to minimize the ohmic voltage drop in the electrolyte. The electrolyte

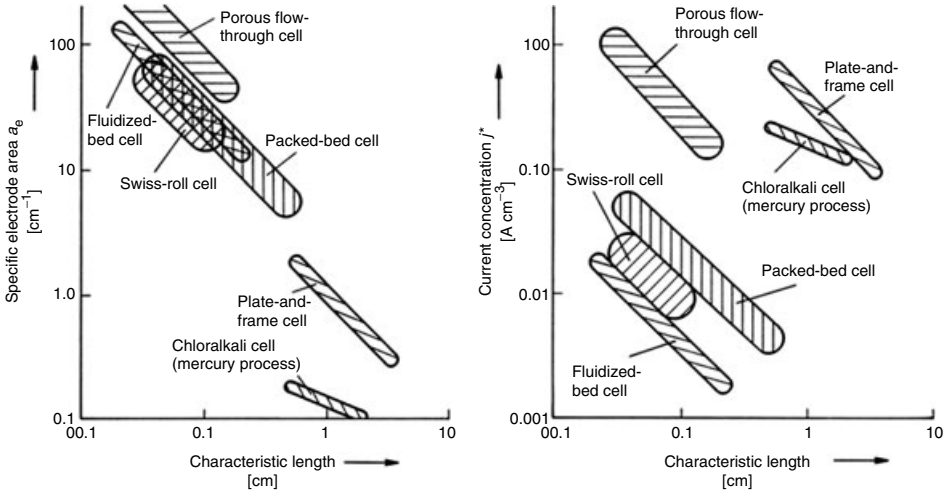


Fig. 5 Specific electrode area a_e and volumetric current concentration j^* as a function of the characteristic length of different electrochemical cells [13].

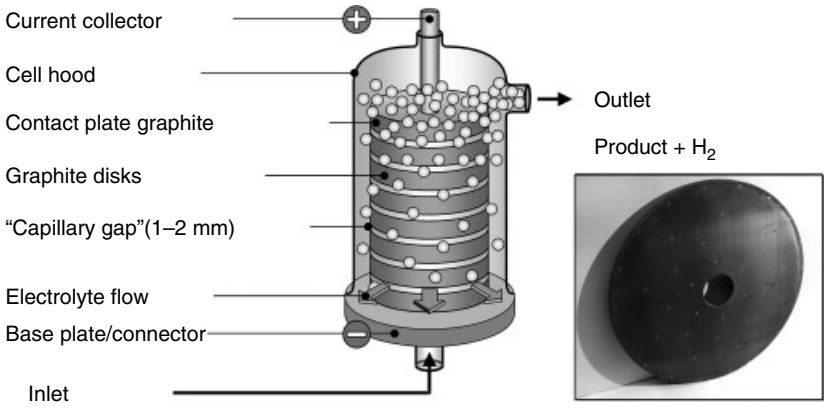


Fig. 6 Scheme of the capillary gap cell of BASF AG [28].



Fig. 7 Cell room with capillary gap electrolyzers [28] (photo: BASF AG).

enters via a central channel and is radially distributed between the circular disk electrodes. In the bipolar configuration, the disks are isolated from each other and connected at the end plates to the voltage supply. To avoid reconversion of the product at the counterelectrodes in the undivided cells, the disk geometry and flow velocity must be carefully tuned to ensure that the diffusion layer thickness, which increases along the radius, does not reach the counterelectrode. Although the scale-up of the stack height and number of disks is simple, the outer disk radius cannot be enlarged. A number of electroorganic processes with methoxylation reactions for production of, for example, anisaldehyde and *tert*-butylbenzaldehyde, were recently developed at BASF [90, 91] using a somewhat modified cell with

graphite felt electrodes. They were also successful in developing ''paired electrolysis'' with the simultaneous formation of valuable products at the anode and at the cathode. A methoxylation at the anode was combined, for example, with the reduction of phthalic acid dimethylester to phthalide, a valuable precursor in pest management chemistry [92].

$$Anode: \text{R-C}_6\text{H}_4\text{-CH}_3 - 4\text{e}^- + 2\ \text{CH}_3\text{OH} \longrightarrow \text{R-C}_6\text{H}_4\text{-CH(OCH}_3)_2 + 4\ \text{H}^+$$

$$Cathode: \text{C}_6\text{H}_4(\text{COOCH}_3)_2 + 4\text{e}^- + 4\ \text{H}^+ \longrightarrow \text{C}_8\text{H}_8\text{O}_2 + 2\ \text{MeOH}$$

Another cell design strictly directed toward minimization of the ohmic drop in the electrolyte, especially if gases are developed at the electrodes, is the zero-gap cell, shown in Fig. 8 [17, 93, 94]. The perforated electrodes are pressed directly onto the diaphragm by the current collectors providing optimum contact across the whole electrode area. However, uneven current distribution and voltage drop in the electrolyte as well as in the diaphragm cannot be avoided, as illustrated in Fig. 8. Zero-gap cells have been successfully used for alkaline water electrolysis [93, 94].

Another design that is used in chloralkali electrolysis, water electrolysis, and electro-organic synthesis [95–97] is the solid polymer electrolyte (SPE) cell, where an ion exchanger membrane, for example, Nafion®, serves as the electrolyte, Fig. 9. The microporous catalytic reaction layers are pressed directly onto the membrane with porous current collectors allowing transport of dissolved reactants and gaseous products into and out of the reaction layer.

The advantages of SPE technology are as follows: electro-organic synthesis without an additional supporting electrolyte, reduced energy demand for separation and recycling processes, and elimination of side reactions with the electrolyte at moderate reaction conditions with ease of process control.

The structure of the SPE electrolysis cell is similar to that of the membrane

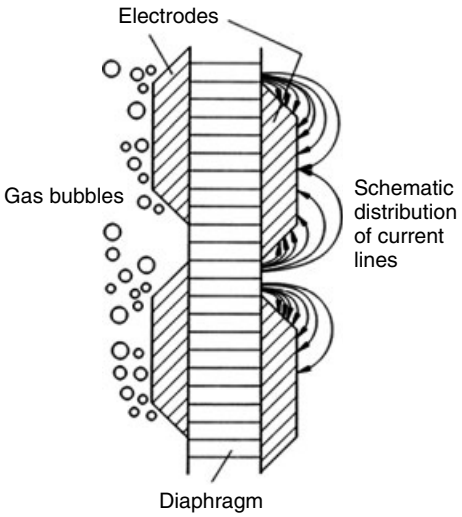


Fig. 8 Gas evolution and current distribution at perforated zero-gap electrodes [13].

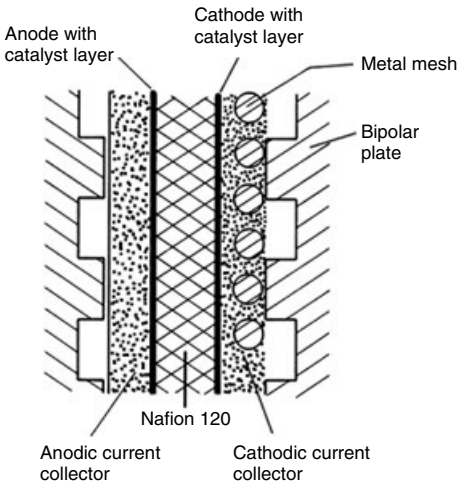


Fig. 9 Solid polymer electrolyte (SPE) cell with cation exchanger membrane Nafion® [13].

electrode assembly in polymer electrolyte membrane (PEM) fuel cell technology, which is described in detail in Chapter 8.

References

1. C. Wagner, The scope of electrochemical engineering in *Advances in Electrochemistry and Electrochemical Engineering* (Eds.: P. Delahaye, C. W. Tobias), Wiley-Interscience, New York, London, Sydney, 1962, p. 1, Vol. 2.
2. C. L. Mantel, *Electrochemical Engineering*, 4th ed., MacGraw-Hill, New York, 1960.
3. O. Levenspiel, *Chemical Reaction Engineering*, 2nd ed., wiley, New York, 1972.
4. J. S. Newman, *Electrochemical Systems*, Prentice-Hall, Englewood Cliffs, 1973.
5. D. J. Pickett, *Electrochemical Reactor Design*, 2nd ed., Elsevier, Amsterdam, 1979.
6. F. Beck, *Elektroorganische Chemie*, VCH, Weinheim, 1974.
7. J. O. M. Bockris, B. E. Conway, E. Yeager, *Comprehensive Treatise of Electrochemistry*, Plenum Press, New York, Vol. 2, Electrochemical Processing, 1981, Vol. 3, Electrochemical Energy Conversion and Storage, 1981, Vol. 4, Electrochemical Material Science, 1981, Vol. 6, Electronics: Transport, 1983.
8. F. Coeuret, A. Storck, *Elements de Genie Electrochimique*, Tecdoc, Paris, 1984.
9. F. Hine, *Electrode Processes and Electrochemical Engineering*, Plenum Press, New York, 1985.
10. T. Z. Fahidy, *Principles of Electrochemical Reactor Analysis*, Elsevier, Amsterdam, 1985.
11. I. Rouzar, K. Micka, A. Kimla, *Electrochemical Engineering*, Elsevier, Amsterdam, 1986, Vol. 1 and 2.
12. E. Heitz, G. Kreysa, *Principles of Electrochemical Engineering*, VCH, Weinheim, 1986.
13. F. Beck, H. Goldacker, G. Kreysa et al., Electrochemistry in *Ullmann's Encyclopedia of Industrial Chemistry*, VCH, Weinheim, 1987, pp. 183–254, Vol. A9.
14. K. Scott, *Electrochemical Reaction Engineering*, Academic Press, London, 1991.
15. D. Pletcher, F. C. Walsh, *Industrial Electrochemistry*, Chapman and Hall, London, 1993.
16. F. Goodridge, K. Scott, *Electrochemical Process Engineering*, Plenum Press, New York, 1995.
17. H. Wendt, G. Kreysa, *Electrochemical Engineering*, Springer, Berlin, 1999.
18. V. M. Schmidt, *Elektrochemische Verfahrenstechnik*, Wiley-VCH, Weinheim, 2003.
19. J. O. M. Bockris, *Electrochemistry of Cleaner Environments*, Plenum, New York, 1972.
20. G. Kreysa, Reactor design for electrochemical water treatment in *Process Technologies for Water Treatment* (Ed.: S. Stucki), Plenum Publishing Corporation, New York, 1988.
21. D. Pletcher, N. L. Weinberg, The Green Potential of Electrochemistry, Chemical

Engineering (London), 1992, Aug. 98–103 (Part 1), Nov. 132–141 (Part 2).
22. K. Rajeshwar, J. G. Ibanez, G. M. Swain, *J. Appl. Electrochem.* **1994**, *24*, 1077.
23. C. A. C. Sequeira, *Environmentally Oriented Electrochemistry*, Elsevier, Amsterdam, 1994.
24. S. Trasatti, Electrochemistry and the environment: the role of electrocatalysis, *Int. J. Hydrogen Energy* **1995**, *20*, 835.
25. K. Rajeshwar, J. G. Ibanez, *Fundamentals and Application in Pollution Abatement*, Academic Press, San Diego, 1997.
26. D. Simonson, Electrochemistry for a cleaner environment, *Chem. Soc. Rev.* **1997**, *26*, 181.
27. K. Jüttner, U. Galla, H. Schmieder, Electrochemical approaches to environmental problems in the process industry, *Electrochim. Acta* **2000**, *45*, 2575.
28. D. Hoormann, J. Jörissen, H. Pütter, *Chem.-Ing.-Tech.* **2005**, *77*, 1363–1376.
29. P. Delahay, *Double Layer and Electrode Kinetics*, Wiley Interscience, New York, 1965.
30. K. J. Vetter, *Electrochemical Kinetics*, Academic Press, New York, 1967.
31. P. W. Atkins, *Physical Chemistry*, 3rd ed., Oxford University Press, Oxford, 1986.
32. J. Koryta, J. Dvorak, *Principles of Electrochemistry*, Wiley, Chichester, 1987.
33. D. D. Macdonald, *Transient Techniques in Electrochemistry*, Plenum Press, New York, 1977.
34. J. Lipkowsky, P. Ross, *Frontiers of Electrochemistry*, Wiley, New York, Vol. 1: Adsorption of Molecules at Metal Electrodes, 1992, Vol. 2: Structure of Electrified Interfaces, 1993.
35. C. M. A. Brett, A. M. O. Brett, *Electrochemistry, Principles, Methods and Applications*, Oxford University Press, Oxford, 1993.
36. E. Gileadi, J. Wang, *Electrode Kinetics*, Wiley, New York, 1994.
37. K. B. Oldham, J. C. Myland, *Fundamentals of Electrochemical Science*, Academic Press, San Diego, CA, 1994.
38. A. Mozik, W. Schmickler, F. Willig, *Surface Electron Transfer Processes*, VCH, Weinheim, 1995.
39. W. Schmickler, *Interfacial Electrochemistry*, Oxford University Press, Oxford, 1996.
40. E. B. Budevski, G. T. Staikov, W. J. Lorenz, *Electrochemical Phase Formation and Growth*, Wiley, New York, 1996.
41. A. Hamnett, C. H. Hamann, W. Vielstich, *Electrochemistry*, VCH, Weinheim, 1997.
42. J. O. M. Bockris, A. K. N. Reddy, M. Gamboa-Aldeco, *Modern Electrochemistry 2A – Fundamentals of Electronics*, Kluwer, New York, 1998.
43. M. Paunovic, *Fundamentals in Electrochemical Deposition*, Wiley, New York, 1998.
44. W. J. Lorenz, W. Plieth, *Electrochemical Nanotechnology*, Wiley-VCH, Weinheim, 1998.
45. J. Grimshaw, *Electrochemical Reactions and Mechanisms in Organic Chemistry*, Elsevier, Amsterdam, 2000.
46. R. Memming, *Semiconductor Electrochemistry*, Wiley-VCH, Weinheim, 2000.
47. J. Bard, L. R. Faulkner, *Electrochemical Methods*, 2nd ed., Wiley, New York, 2001.
48. J. W. Schultze, G. Staikov, *Scaling Down in Electrochemistry: Electrochemical Micro- and Nanosystem Technology*, Elsevier, Amsterdam, 2001.
49. H. Kaesche, *Corrosion of Metals, Physicochemical Principles and Current Problems*, Springer, Berlin, 2003.
50. H. H. Girault, *Analytical and Physical Electrochemistry*, EPFI Press, Marcel Dekker, 2004.
51. *International Union of Pure and Applied Chemistry: Größen, Einheiten und Symbole in der Physikalischen Chemie* (Ed.: K.-H. Homann), VCH, Weinheim, 1996.
52. G. Gritzner, G. Kreysa, Nomenclature, symbols and definitions in electrochemical engineering, *J. Electroanal. Chem.* **1993**, *360*, 351.
53. Handbook of Chemistry and Physics, Robert and Weast, 1977.
54. I. Barin, O. Knacke, O. Kubaschewski, *Thermodynamical Properties of Inorganic Substances*, Springer, Berlin, 1973.
55. R. E. W. Janson, R. J. Marshall, *Chem. Eng.* **1976**, *315*, 769.
56. R. E. W. Janson, N. R. Tomov, *Chem. Eng.* **1977**, *327*, 867.
57. C. L. Lopez-Cacicedo, *Inst. Chem. Eng. Symp. Ser.* **1975**, *42*, 29.
58. D. R. Gabe, *J. Appl. Electrochem.* **1974**, *4*, 91.
59. F. S. Holland, *Chem. Ind.* (London), **1978**, *13*, 453.
60. L. J. Ricci, *Chem. Eng.* **1975**, *82*(27), 29.
61. R. Kammel, H.-W. Lieber, *Galvanotechnik* **1977**, *68*, 710.
62. W. Götzelmann, *Galvanotechnik* **1977**, *68*, 789.
63. D. Bruhn, W. Dietz, K.-J. Müller, C. Reynvaan, EPA 86109265.8, 1986.

64. A. Storck, P. M. Robertson, N. Ibl, *Electrochim. Acta* **1979**, *24*, 373.
65. P. M. Robertson, F. Schwager, N. Ibl, *J. Electroanal. Chem.* **1975**, *65*, 883.
66. K. B. Keating, J. M. Williams, *Res. Rec. Conserv.* **1976**, *2*, 39.
67. F. Beck, H. Guthke, *Chem.-Ing.-Tech.* **1969**, *41*, 943.
68. D. N. Bennion, J. Newman, *J. Appl. Electrochem.* **1972**, *2*, 113.
69. G. Kreysa, *Chem.-Ing.-Tech.* **1983**, *55*, 23.
70. G. Kreysa, C. Reynvaan, *J. Appl. Electrochem.* **1982**, *12*, 241.
71. G. Kreysa, Elektrochem. Zelle, DE 26 22 497, 1976.
72. J. R. Backhurst, J. M. Coulson, F. Goodridge, et al., *J. Electrochem. Soc.* **1969**, *116*, 1600.
73. J. R. Backhurst, M. Fleischmann, F. Goodridge, R. E. Plimley, GB Pat, 1 194 181, 1970.
74. H. Scharf, DE 22 27 084, 1972.
75. C. Raats, H. Boon, W. Eveleens, *Erzmetall* **1977**, *30*, 365.
76. G. van der Heiden, C. Raats, H. Boon, *Chem. Ind.* (London), **1978**, *13*, 465.
77. R. Kammel, H.-W. Lieber, *Galvanotechnik* **1978**, *69*, 687.
78. W. Götzelmann, *Galvanotechnik* **1979**, *70*, 596.
79. D. R. Gabe, F. C. Walsh, *Trans. IchemE* **1990**, *68B*, 107.
80. G. Kreysa, *Chem.-Ing.-Tech.* **1978**, *50*, 332.
81. G. Kreysa, *Metalloberfäche* **1980**, *34*, 494.
82. G. A. Carlson, E. E. Estep, D. Jacqueau, *Chem.-Ing.-Tech.* **1973**, *45*, 217.
83. R. S. Wenger, D. N. Bennion, *J. Appl. Electrochem.* **1976**, *6*, 385.
84. J. van Zee, J. Newman, *J. Electrochem. Soc.* **1977**, *124*, 706.
85. D. S. Flett, *Chem. Ind.* (London), **1971**, *10*, 300.
86. D. S. Flett, *Chem. Ind.* (London), **1972**, *14*, 983.
87. K. Scott, *J. Appl. Electrochem.* **1988**, *18*, 504.
88. H. Bergmann, H. Hertwig, F. Nieber, *Chem. Eng. Proc.* **1992**, *31*, 195.
89. J. Jörissen, *Practical Aspects of Preparative Scale Electrochemistry, Encyclopedia of Electrochemistry*, Organic Electrochemistry (Ed.: H. Schäfer), Viley-VCH, Weinheim, 2004, Vol 8, Chapter 2.
90. H. Pütter, Industrial electroorganic chemistry in *Hrsg., Organic Electrochemistry* (Eds.: H. Lund, O. Hammerich), 4th ed., Marcel Dekker, New York, 2001, pp. 1259.
91. N. Aust, A. Fischer, T. Freund, U. Griesbach, G. Heyderich, H. Pütter, D. Zöllinger, New electrochemical oxidations – green and sustainable chemistry for valuable intermediates, ISE 55^{th} Annual Meeting, Thessaloniki, 2004.
92. H. Hannebaum, H. Pütter, *Chemie in unserer Zeit* **1999**, *33*, 373–374.
93. J. Fischer, H. Hofmann, G. Luft, et al., *AIChE J.* **1980**, *26*, 794.
94. J. Divisek, H. Schmitz, J. Mergel, *Chem.-Ing.-Tech.* **1980**, *52*, 465.
95. S. Stucki, A. Menth, *Ber. Bunsen-Ges. Phys. Chem.* **1980**, *84*, 1008.
96. J. Jörissen, *Electrochim. Acta* **1996**, *41*, 553–562.
97. J. Jörissen, *J. Appl. Electrochem.* **2003**, *33*, 969–977.

2
Electrocatalysis

A. Katsaounis, S. Brosda, and C.G. Vayenas
University of Patras, Patras, Greece

2.1 Introduction **25**

2.2 Adsorption **28**

2.3 Basic Electrokinetic Principles and Experimental Techniques **34**
2.3.1 Electrokinetic and Mass-transport Control 34
2.3.2 The Measurement of IR-free Electrocatalyst Overpotential 36
2.3.3 Cyclic Voltammetry 38
2.3.4 AC-impedance Spectroscopy 43
2.3.5 Normalization of Active Surface 47

2.4 Theory of Electrocatalysis **48**
2.4.1 Overpotentials and the Butler–Volmer Equation 48
2.4.2 Volcano Plots 51
2.4.3 Transition State Approaches 53
2.4.4 Quantum Mechanics–based Models 54
2.4.5 Promotional Effects 57

2.5 The Hydrogen Oxidation/Evolution Reaction **58**

2.6 Oxygen Reduction/Evolution Reaction **60**

2.7 Oxidation of Small Organic Molecules **63**

2.8 Other Important Electrocatalytic Systems **66**

2.9 Electrochemical Promotion of Catalytic Reactions **69**

2.10 Promotional Rules in Catalytic and Electrocatalytic Systems **76**

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

2.11 Concluding Remarks **80**
References 80

2.1 Introduction

The term *electrocatalysis* was first introduced by Grubb in 1963 [1] in connection with the anodic and cathodic charge-transfer reactions in fuel cells. However, the first systematic experimental investigations of various electrocatalysts had been carried out in the 1920s by Bowden and Rideal [2] and the concept and first interpretation of electrocatalysis had been introduced in the 1930s by Horiuti and Polanyi [3]. Their theory and its extensions and improvements have been analyzed lucidly by Bockris [4–6]. Recent reviews of progress in electrocatalysis can be found in more general [6–8] or specialized books [9–11].

An electrocatalyst is an electrode that can accelerate the approach to equilibrium of an electrochemical process, which involves *a net charge transfer*, without itself being consumed in the process.

This definition is in analogy to that of a catalyst which is a substance (quite often a solid) which can accelerate, without being consumed, the approach to equilibrium of a chemical process in a gaseous or liquid mixture.

In heterogeneous catalysis, the specific activity of a catalyst is usually expressed as turnover frequency (TOF) (s^{-1}) which expresses the number of catalytic turnovers per surface catalyst atom per second. Typical turnover frequencies of industrial catalytic processes are 10^{-3} to $10^{2}\ s^{-1}$. Slower TOFs are economically uninteresting and larger ones are usually not realizable in a practical environment because of diffusional limitations.

In electrocatalysis, the activity of different electrocatalysts is usually expressed via the exchange current I_o, and the specific activity, via the exchange current density, i_o (A cm^{-2}), still often computed on the basis of the superficial electrode surface area. Only when the current is normalized using the true surface area of the electrode–electrolyte interface, the comparison between different electrocatalysts is truly meaningful. The determination of the true surface area of porous electrodes is discussed in Sect. 2.3.5.

For a Pt(111) surface, with a surface density of 1.5×10^{15} atoms cm^{-2}, the current density corresponding to a TOF of $1\ s^{-1}$ is 0.18 mA cm^{-2} for a one-electron charge-transfer process. Such exchange current densities based on the real electrocatalyst surface area are quite typical for a decent electrocatalyst for H_2 evolution or oxidation. Thus, one may conclude that the order of magnitude of the TOFs of catalytic and electrocatalytic reactions are quite similar. In the latter

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

case, of course, the TOF can be very significantly modified by the electrode potential.

Different electrodes can have dramatically different electrocatalytic activity. An example is given in Fig. 1 which shows current–potential curves obtained with different electrodes for the same charge-transfer reaction, that is, H_2 evolution. One observes that at the same overpotential, the current density, i, obtained with different electrocatalysts and, thus, also the specific rate of H_2 evolution i/F, where F is the Faraday constant (96 450 C mol^{-1}), varies by many orders of magnitude. In this figure, the current has been normalized using the superficial electrode surface area and this is the main reason for the very large differences that are observed between ''bright'' Pt and ''spongy'' Pt (Fig. 1).

Horiuti and Polanyi [3] were the first to provide an interpretation of electrocatalysis (Fig. 2). The potential energy of a proton moving toward the electrode exhibits a maximum at the intersection of the potential energy curve of the proton in solution and that of a H atom adsorbed on the electrode. A good electrocatalyst is that in which this maximum is small. Upon varying the electrode potential, the energy of adsorption varies by $\Delta\Delta H_{ads}$ and thus the activation barrier changes by $\beta\Delta\Delta H_{ads}$ where the parameter β (typically of the order of 0.5–1) dictates, to a large extent, the slope of the Tafel diagrams of Fig. 1.

Although, during the last 50 years, significantly more detailed classical or quantum mechanical treatments have been formulated to describe and predict the classical Tafel behavior of electrocatalysis (Fig. 1) [12–22], the basic ideas of the Horiuti and Polanyi interpretation remain fruitful. In fact, the linear variation in adsorption enthalpies with potential is well established in high-temperature solid electrolyte electrochemistry [23, 24] (Fig. 3), where enthalpies and entropies of adsorption are quite straightforward to measure using standard surface science techniques, such as temperature-programmed desorption (TPD). This is similar with the observed linear variation of the adsorption enthalpies of alkalis on metal surfaces with the adsorption-induced change in catalyst work function [23, 25] (Fig. 4). It is worth noting in Figs 3 and 4 that the slope of the

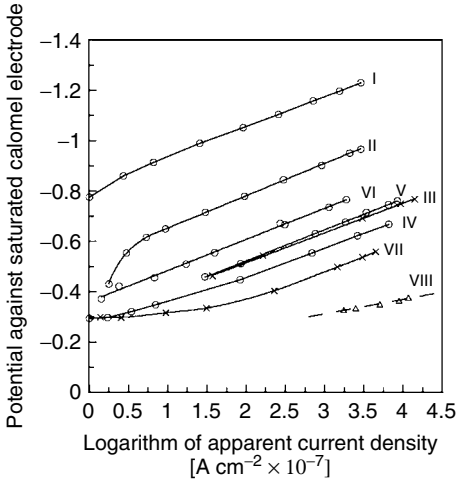


Fig. 1 Some of the first published comparative current–potential (Tafel) plots obtained by Bowden and Rideal [2, 4]. I. Mercury. II. Platinized mercury, 1/100 covered. III. Platinized mercury, thin film of Pt. IV. Etched silver, new. V. Etched silver, old. VI. Polished silver. VII. Bright platinum. VIII. Spongy platinum. (In Curve VIII the current density is expressed in amperes $\times 10^{-6}$ in order to bring it on the figure. It should be displaced two divisions to the right.) (Reprinted with permission from Ref. 2, Copyright 1928 by The Royal Society).

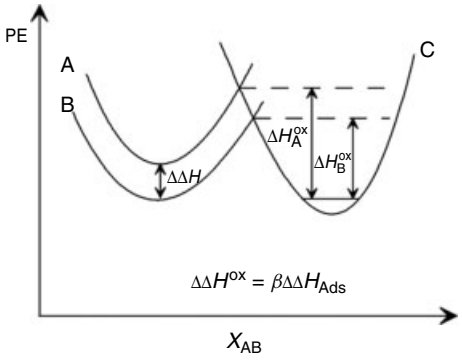


Fig. 2 The effect of the variation of the $\Delta\Delta H$ on the potential energy–distance diagram. (Reprinted with permission from Ref. 4, Copyright 1994 by Elsevier).

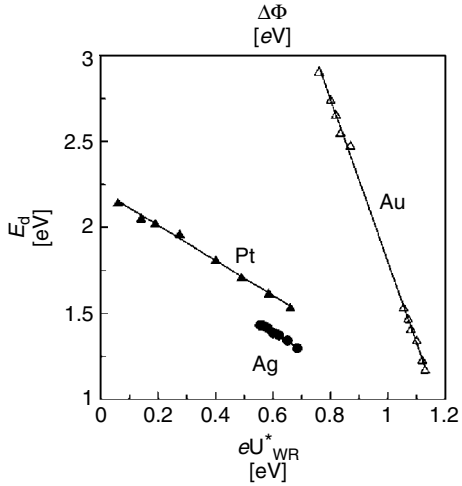


Fig. 3 Effect of catalyst potential on the oxygen desorption activation energy, E_d, calculated from the Redhead analysis of O_2 TPD spectra for Pt, Ag, and Au electrodes deposited on YSZ. (Reprinted with permission from Ref. 24, Copyright 1997 by Springer).

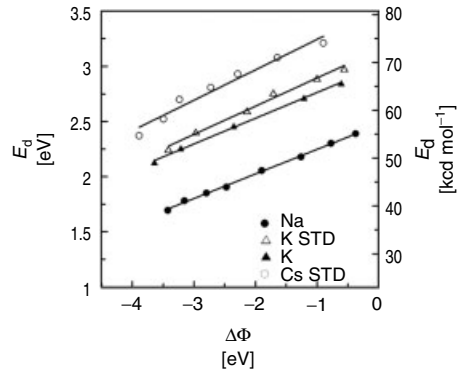


Fig. 4 Effect of work function change $\Delta\Phi$ on the alkali desorption energy [23, 25]. (Reprinted with permission from [25], Copyright 1987 by Elsevier).

straight lines are of the order −1 for atomic oxygen and +1 for atomic potassium, an observation supported by rigorous quantum mechanical cluster treatments [23, 26]. This linear variation in adsorption enthalpies with potential is also responsible for the effect of electrochemical promotion on non-Faradaic electrochemical modification of catalytic activity (NEMCA) effect where electrocatalysis is used to generate promoters for catalytic (no net charge transfer) reactions in situ. It has been shown [26] that, in most cases, these linear on near-linear variations in adsorption enthalpies with potential reflect the fact that adsorption involves formation of a chemical (chemisorptive) bond, between the adsorbate and the electrode, in the presence of a very strong electric field (typically 10^{10} V m^{-1} or 1 V $Å^{-1}$) created in the double layer formed at the electrode–liquid or electrode–gas interface owing to the presence of the adsorbate and other weakly or strongly coadsorbed species. This electric field varies linearly with potential (it vanishes at the potential of zero charge, pzc) [23, 27] and its interaction with the adsorbate dipoles is the main cause for the observed linear variation in adsorption enthalpies with potential.

During the last 15 years, the urgent need for developing practical fuel cell devices has led to an unprecedented interest in electrocatalysis for the development of efficient anodes and cathodes both in aqueous [9] and in high-temperature [28] electrochemistry.

2.2 Adsorption

The role of adsorption in electrocatalysis is as important as it is in heterogeneous catalysis. In aqueous electrochemistry, the presence of H_2O makes the study of adsorption more complex than in solid-state electrochemistry or in heterogeneous catalysis. The basic principles of the subject are analyzed lucidly by Gileadi [29].

The simplest isotherm in electrocatalysis is the electrochemical Langmuir isotherm. Considering, for example, the adsorption of chlorine atoms on a metal electrode, M, one can write:

$$Cl^- + M \rightleftharpoons M - Cl + e \qquad (1)$$

and the electrochemical Langmuir isotherm is

$$\theta/(1-\theta) = Kc\exp[FU/RT] \qquad (2)$$

where θ is the coverage of Cl on the metal surface, U is the electrode potential, c is the concentration of Cl^- ions in the solution, and K is an adsorption equilibrium constant.

If the adsorption is dissociative, for example

$$2M + H_2 \rightleftharpoons 2H^+ + 2e^- + 2M$$
$$\rightleftharpoons 2M - H \qquad (3)$$

then the Langmuir isotherm can be expressed as

$$\theta^2/(1-\theta)^2 = Kp_{H_2}$$
$$= Kc^2\exp[-FU/RT] \qquad (4)$$

where c is the proton concentration.

If the adsorption involves a large molecule which occupies n sites on the surface, then Eq. (2) can be written in the form

$$\theta/(1-\theta)^n = Kc\exp[FU/RT] \qquad (5)$$

where the sign of the potential in (5), as well as in Eqs (2) and (4), is dictated by whether an electron is taken from or given to the electrocatalyst during the adsorption

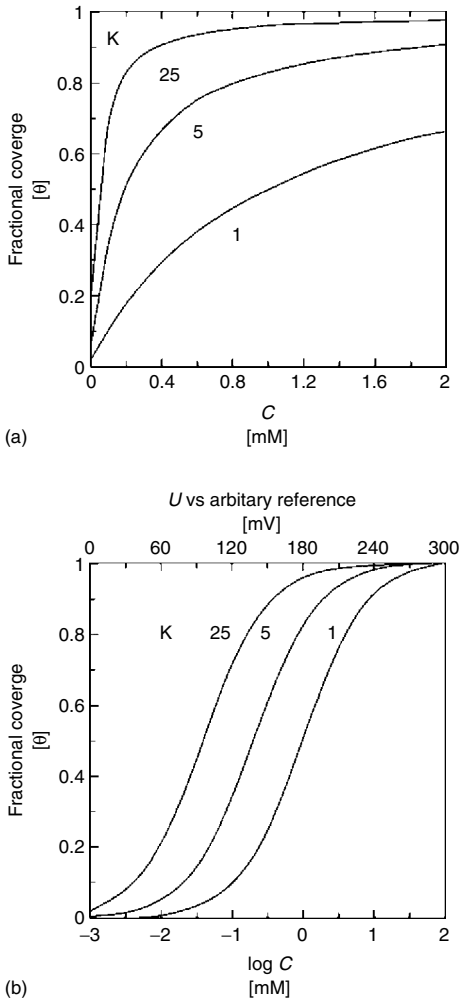


Fig. 5 Electrochemical Langmuir isotherm. Plots of fractional coverage θ versus (a) concentration and (b) potential or log c, for three values of the equilibrium constant $K = 1$, 5, and 25. (Reprinted with permission from Ref. 29, Copyright 1993 by Wiley-VCH).

process. Thus, the sign is positive when the adsorbate starts as an anion and negative when the adsorbate starts as a cation.

Figure 5 shows plots of Eq. (2) for various values of K. In Fig. 5(a), the coverage θ is plotted versus c, while in Fig. 5(b) it is plotted versus log c. In the latter case, increasing K causes a parallel shift in the θ versus log c plot. A change in the potential equal to $2.3RT/F$ has the same effect on θ as a change of a factor of 10 in the ion concentration.

The assumptions of the electrochemical Langmuir isotherm are the same as those for the Langmuir isotherm in heterogeneous catalysis, that is, surface homogeneity and negligible lateral adsorbate interactions. This leads to a K value which does not depend on θ and, thus, to a standard free energy of adsorption, ΔG°, which also does not depend on θ. Note that in the context of the Langmuir or electrochemical Langmuir adsorption isotherm the standard state of the adsorbate is at $\theta = 1/2$.

The first refinement of the electrochemical Langmuir isotherm is the Frumkin isotherm, which is obtained by assuming a linear dependence of ΔG° on θ, that is,

$$\Delta G^\circ = \Delta G^\circ_o + r\theta \qquad (6)$$

where the parameter r (typically $|r| = 20-100\ \mathrm{kJ\ mol^{-1}}$) is a measure of the surface heterogeneity and, usually more importantly, a measure of the strength of lateral adsorbate–adsorbate interactions, which can be mediated either via the aqueous or gas phase or via the metal electrode itself. A positive r implies repulsive adsorbate–adsorbate interactions and a negative r implies attractive adsorbate interactions.

The resulting isotherm

$$\theta/(1-\theta)\exp(r\theta/RT) = K_o c \exp(FU/RT) \qquad (7)$$

is known as the *Frumkin isotherm*.

Figure 6 shows the dependence of θ on potential for a fixed value of c. The plots of θ versus c at fixed potential are qualitatively similar. For $r/RT < -4$, one obtains steady-state multiplicity, that is, for fixed c there are three steady-state coverage values, two of which (the lowest and the highest) are stable, and usually very different. The high-coverage state corresponds to surface "condensation", or island formation, due to the attractive lateral interactions which lead to an ordered two-dimensional adsorbate surface structure.

The Frumkin isotherm is mathematically equivalent to the effective double-layer (EDL) isotherm recently introduced in heterogeneous catalysis [23, 27] and which accounts for the existence of an overall neutral double layer at the metal–gas interface. This type of isotherm is very useful in describing the effect of electrochemical promotion of catalysis (EPOC or NEMCA effect), which is further discussed in the following text, and refers to the pronounced and strongly non-Faradaic change in catalytic activity observed upon electrochemical supply or removal of ions to or from a catalyst surface. This effect has been studied in much more detail in solid-state electrochemistry, utilizing Y_2O_3-stabilized ZrO_2 (YSZ), (an O^{2-} conductor) or β''-Al_2O_3 (a Na^+ conductor) [23], than in aqueous systems [23, 30, 31], where, on the other hand, the NEMCA effect was recently shown to play

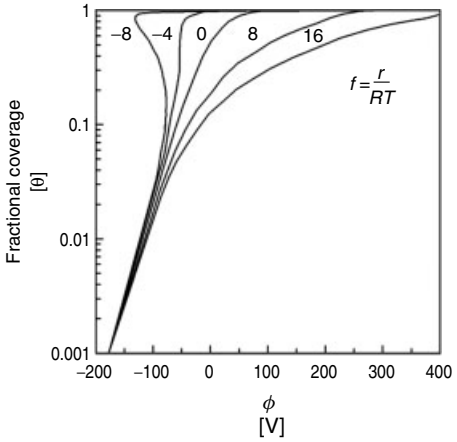


Fig. 6 Frumkin isotherm, plotted as log θ versus U, for different values of $f = r/RT$. Negative values of f correspond to a lateral attraction interaction (i.e., to an increase of $|\Delta G^\circ_\theta|$ with coverage). The standard potential for adsorption, U^o_θ, is defined as the potential where $[\theta/(1-\theta)] = 0$ for $r = 0$ and unit concentration in the solution. (Reprinted with permission from Ref. 29, Copyright 1993 by Wiley-VCH).

a significant role in the operation of CO-poisoned low-temperature polymer electrolyte membrane (PEM) fuel cells [32].

In aqueous electrochemistry, two more factors are important to consider when studying adsorption or using adsorption isotherms. The first is the presence of H_2O.

The state-of-the-art knowledge of water adsorption in ultrahigh vacuum (UHV), up to the late 1980s has been reviewed in detail by Thiel and Madey [33] and in a very lucid and concise form by Leiva [34]. Water bonding occurs with charge transfer to the surface, lowering the work function of the metal significantly, while the internal molecular structure and vibrational properties of water are only slightly perturbed. The adsorption energy of water varies between 0.4 and 0.7 eV. Thus, water is a weakly chemisorbed species when compared to other adsorbates such as CO. The low adsorption energy implies that adsorbed water molecules tend to make hydrogen bonds with each other, even at low coverages. Thus, it is experimentally very difficult to study isolated adsorbed molecules. Water clustering is practically unavoidable. The water structure of the first two monolayers, established from diffraction experiments over hexagonal surfaces, is shown in Fig. 7. This structure consists of an icelike hydrogen-bonded network in which the first layer is bound to the metallic surface through oxygen atoms and the second layer is held to the first layer by hydrogen bonding. The stability of this structure decreases with increasing lattice mismatch between ice and the underlying surface.

The structure of water at the metal–solution interface is further complicated by the presence of an external electric field. One may distinguish three different cases: the surface at the pzc; the surface at a potential more positive than the pzc, thus having a positive charge; and the surface at more negative potentials than the pzc, thus having a negative charge. Owing to the strong dipole moment of water and the high electric fields present at the interface, one expects that the structure of the water–metal interface in the region close to the metal will exhibit significant differences between these three cases. Indeed, this is the case as shown in Fig. 8 which presents some classical X-ray diffraction experiments carried out by Toney's group [34, 35] and shows the oxygen distribution function in the neighborhood of a Ag(111) electrode at a potential more negative (Fig. 8a) and more positive (Fig. 8b) than the pzc for a Ag(111) surface. These distribution functions represent the local increase or decrease in the oxygen density at a certain distance from the surface with respect to the bulk average value. One observes from the three peaks,

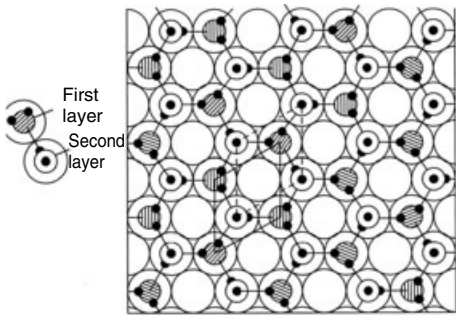


Fig. 7 Schematic diagram of an adsorbed water bilayer over an hexagonal surface. Large empty circles denote surface atoms and small black-filled circles represent hydrogen atoms in water molecules. The remaining circles denote oxygen atoms belonging to adsorbed water molecules. The $(\sqrt{3}\times\sqrt{3})R30^{\circ}$ unit cell is outlined. (Reprinted with permission from Ref. 33, Copyright 1987 by Elsevier).

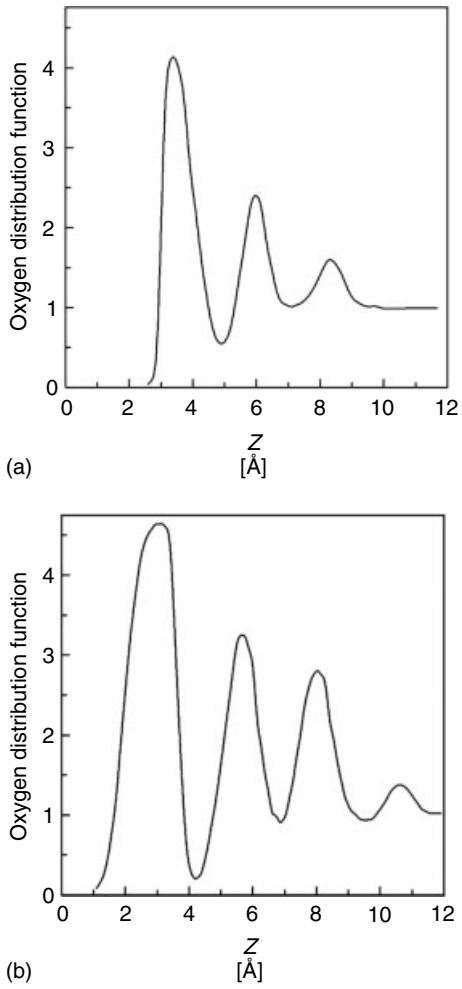


Fig. 8 Oxygen distribution functions in the direction parallel to a Ag(111) electrode as obtained from X-ray diffraction experiments [35]. (a,b) Potentials of −0.23 and +0.52 V with respect to the potential of zero charge respectively. $Z = 0$ corresponds to the center of the first Ag plane. Density is referred to the bulk density of water. (Reprinted with permission from Ref. 35, Copyright 1995 by Elsevier).

next to the surface in both graphs, that the electrode surface produces a layering of the water molecules that extends up to three layers toward the bulk. The differences between positively and negatively charged surfaces are pronounced; at a positive surface (Fig. 8b), charged water molecules are attracted toward the electrode and the density of the layers is increased with respect to the negatively charged surface (Fig. 8a). Additional useful information about the water structure at the metal–solution interface comes from the use of infrared spectroscopic techniques. Both the intensity and the frequency of surface water IR bands change when going from positively to negatively charged surfaces [36]. In Fig. 9, the relative intensity of the HOH bending band, obtained from surface-enhanced infrared absorption spectroscopy (SEIRAS) of a Au(111) surface in contact with a perchloric acid solution, is shown as a function of electrode potential. The observed changes in intensity are due to changes in the orientation of the water molecules at the interface. At the pzc, the water molecules are nearly parallel to the surface and change from oxygen-up to oxygen-down in going from negative to positive charges. The structure at positive potentials is interpreted to be icelike as in Fig. 7. This structure is distorted at potentials over 0.8 V owing to perchlorate anion adsorption. Another example of the abrupt change in interfacial water structure near the pzc is presented in Fig. 10 which shows the potential dependence of the in-bending mode of water adsorbed at a Pt(111) electrode. At a potential near 0.4 V, the band center frequency changes abruptly.

A second important factor to be considered when studying adsorption or using adsorption isotherms in liquid solutions is the change in ion concentration in the proximity of the electrode owing to

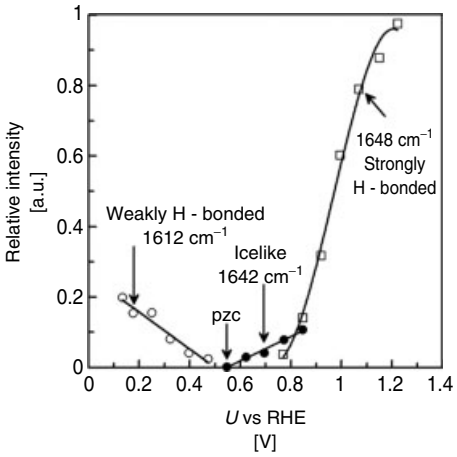


Fig. 9 Plot of the integrated intensities of adsorbed water bending infrared modes as a function of applied potential. Water is adsorbed over a Au(111) electrode. (Reprinted with permission from Ref. 37, Copyright 1996 by the American Chemical Society).

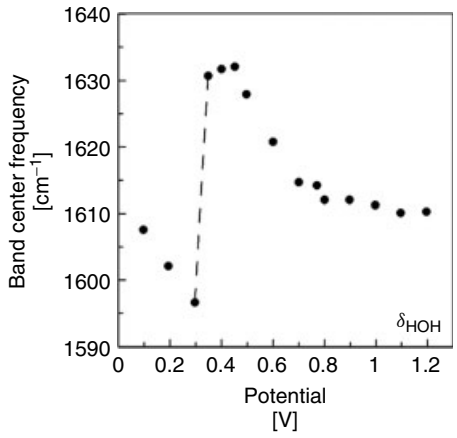


Fig. 10 Potential dependence of the water bending infrared band center frequency as function of applied potential for water adsorbed at a Pt(111) electrode. (Reprinted with permission from Ref. 36, Copyright 1997 by the American Chemical Society).

electrostatic effects. Leiva [34] has used the Gouy-Chapman model and, solving the Poisson–Boltzmann equation

$$\nabla^2\varphi(r) = -\frac{1}{\varepsilon}\sum_i z_i e c_i^{\infty} \exp\left[-\frac{z_i e\varphi(r)}{k_b T}\right] \tag{8}$$

he has obtained Fig. 11(a), which shows the very significant deviation of the ratio c/c_{bulk} from unity when significant potential differences exist at the metal–solution interface. This implies that the surface concentration values, to be used in the electrochemical Langmuir or Frumkin isotherms, may be substantially different from the bulk concentration, c_{bulk}, values. In general, upon adsorption, anions tend to lose part of their solvation layer, which frequently leads to specific adsorption, quite similar to gas-phase adsorption in surface science and heterogeneous catalysis, whereas cations tend to retain their solvation layer. The geometrical centers of specifically adsorbed anions define the inner Helmholz plane (IHP), whereas the centers of cations define the outer Helmholz plane (OHP) as depicted in Fig. 11(b). One may say that the electrode

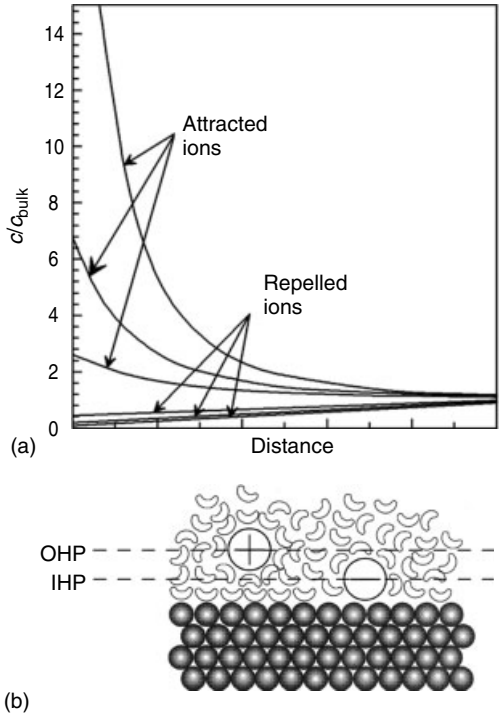


Fig. 11 (a) Effect of distance from electrode on ion concentration. Excess concentration of attracted and repelled ions with respect to bulk concentration in the vicinity of an electrode surface calculated from Gouy-Chapman model [34]. Curves correspond to 0.025, 0.050, and 0.100 V potential differences. (b) Outer Helmholz plane (OHP) and inner Helmholz plane (IHP). Anions adsorbed on the surface have lost part of their solvation layer and are contact adsorbed (IHP); cations remain fully solvated and are water-separated adsorbed (OHP). (Reprinted with permission from Ref. 34, Copyright 2003 by John Wiley & Sons Ltd).

surface and the OHP define the limits of the compact double layer present at the metal–solution interface. This compact double layer creates a strong (10^7 V cm^{-1}) electric field which significantly affects both specific adsorption (chemisorption) and electrocatalysis.

2.3 Basic Electrokinetic Principles and Experimental Techniques

2.3.1 Electrokinetic and Mass-transport Control

Electrocatalysis plays an important role in the performance both of fuel cells, for the production of power, and of electrolytic cells, for the production of chemicals. It also has an important role in electrochemical cogenerative cells where useful chemicals and electrical power are produced simultaneously, for example, via oxidation of NH_3 to HNO_3 or SO_2 to SO_3 [23].

However, the performance of a Galvanic or electrolytic cell is not dictated only by the quality of its anodic and cathodic electrocatalysts but also by the resistance of the electrolyte between the two electrodes, which must be minimized, and by the rates of mass transport of reactants and products to and from the two electrodes.

An illustrative example is shown in Fig. 12 which shows the current–potential curve obtained during electrolysis of a dilute (0.01 M) solution of KI in H_2SO_4 using two Pt electrodes. The minimum potential, U_{rev}, required for the electrolysis reaction is 0.49 V, computed from thermodynamics via $U_{rev} = -\Delta G/nF$,

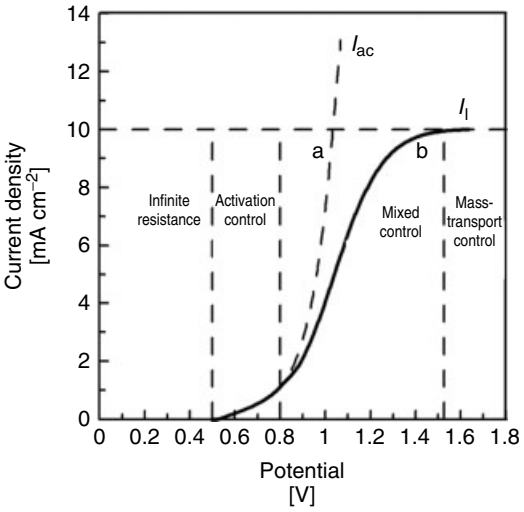


Fig. 12 *I/E* plot for the electrolysis of a dilute (0.01 M) solution of KI in H_2SO_4, employing two Pt electrodes. The minimum potential for current flow is 0.49 V. Line a is purely activation-controlled current I_{ac}, line b is the actual current which will be measured, having a mass-transport-limited value I_l. (Reprinted with permission from Ref. 29, Copyright 1993 by Wiley-VCH).

where ΔG is the Gibbs energy change of the overall reaction, n is the number of electrons transferred, for example, $n = 1$ for the electrolysis of KI, and F is Faraday's constant.

Above this potential difference, the cell current becomes finite and the S-shaped current–potential curve shown in Fig. 12 is obtained. The deviation in potential, $U - U_{rev}$, is the total cell overpotential.

As already shown in Fig. 1, a general feature of electrocatalysis is that the current passing through an electrode–electrolyte interface depends exponentially on overpotential, as described by the Butler–Volmer equation discussed in Sect. 2.4.1, so that $\log i$ versus $\eta(U - U_{rev})$ gives straight lines, termed *Tafel plots* (Fig. 1). On this basis, one would expect an exponential-type dependence of current on overpotential in Fig. 12 (curve labeled I_{ac}). Such a curve would correspond to pure activation control, that is, to infinitely fast mass-transport rates of reactants and products to and from the two electrodes.

However, in any practical cell, these mass-transport rates have a maximum, corresponding to zero concentration of the reactants or very high concentration of the products on the electrode surface, and this maximum, denoted I_l (for diffusion-limited current) dictates the actual shape of the current–potential curve at high overpotentials. Under these conditions, the cell operation is mass transport controlled (Fig. 12).

The total cell activation overpotential is the sum of the activation overpotentials at the anode and cathode, as shown in Fig. 13 for the case of H_2O electrolysis using Pt electrodes in alkaline solutions. The two overpotentials can be separated by the use of a reference electrode. Thus, the use of reference electrodes is essential for the study of electrocatalysis, since in this case one can individually study the dependence of each electrode overpotential on the current and thus assess the electrocatalytic performance of each electrode. The best electrocatalyst, for each charge-transfer reaction is, obviously, the one that minimizes the activation overpotential.

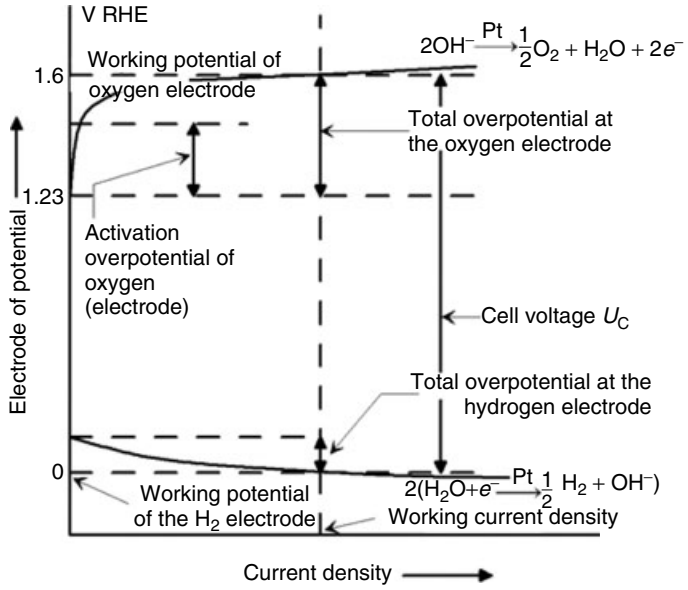


Fig. 13 Current–voltage diagram showing the meaning of activation overpotential for the electrolysis of water using platinum electrodes in alkaline solution. (Reprinted with permission from Ref. 8, Copyright 1998 by Wiley-VCH).

2.3.2 The Measurement of IR-free Electrocatalyst Overpotential

An important aspect of the study of electrocatalysis is the accurate measurement of the overpotential, η, of the electrocatalyst. The overpotential is the deviation of the electrocatalyst potential from its open-circuit ($I = 0$) value. The basic experimental setup for the study of electrocatalysis is the three-electrode system comprising the electrocatalyst under study, termed *working electrode* (W), a *counterelectrode* (C), and a *reference electrode* (R).

When a current I flows between the working electrode (W) and the counterelectrode (C), then the potential difference between them, U_{WC}, deviates from its open-circuit value U^{o}_{WC} (Fig. 14). The electrochemical cell overpotential η_{WC} is then defined from

$$\eta_{WC} = U_{WC} - U^{o}_{WC} \qquad (9)$$

The cell overpotential η_{WC} is the sum of three terms:

$$\eta_{WC} = \eta_W + \eta_C + \eta_{ohmic,WC} \qquad (10)$$

where η_W, η_C are the overpotentials of the catalyst (W) and counter (C) electrodes, respectively, and $\eta_{ohmic,WC}$ is the ohmic overpotential due to the resistance of the electrolyte between the working and counterelectrodes.

The latter equals $I R_{WC}$ where R_{WC} is the ohmic resistance between the working and counterelectrodes. Experimentally, it is rather easy to measure the $\eta_{ohmic,WC}$ term using the current interruption technique as shown in Fig. 14. Upon current interruption, the ohmic overpotential $\eta_{ohmic,WC}$

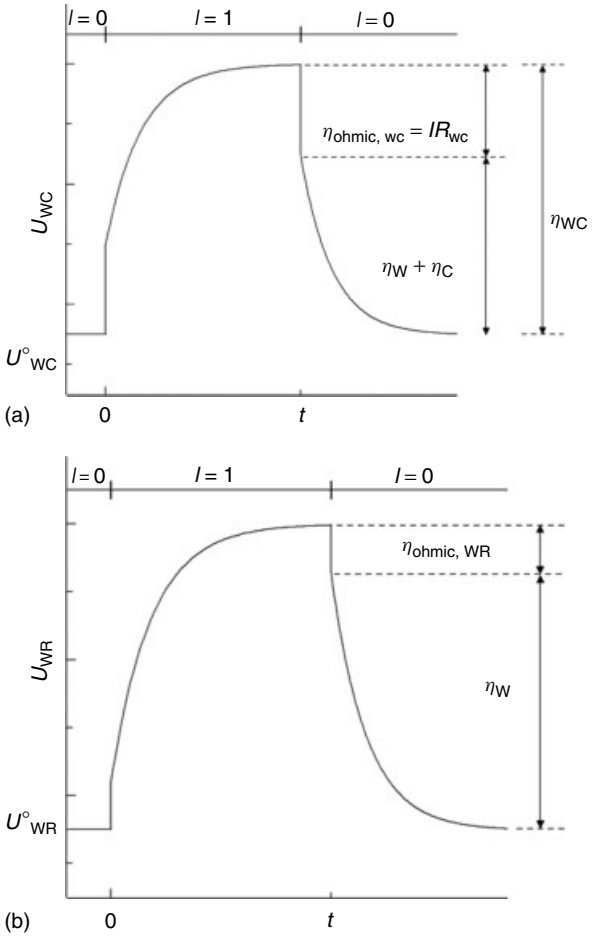


Fig. 14 (a) Use of the current interruption technique to measure the ohmic overpotential, $\eta_{ohmic,WC}$, between the working (W) and counter (C) electrodes. (b) Use of the current interruption technique to measure the uncompensated ohmic drop, $\eta_{ohmic,WR}$, between the working (W) and reference (R) electrodes. (Reprinted with permission from Ref. 23), Copyright 2003 by Kluwer/Plenum Publisher).

vanishes within less than 1 μs and the remaining part of the overpotential which vanishes much slower is $\eta_W + \eta_C$.

It is worth emphasizing that although overpotentials are usually associated with electrode–electrolyte interfaces, in reality they refer to, and are measured as, deviations of the potential of the electrodes only. Thus, the concept of overpotential must be associated with *an electrode* and not with an electrode–electrolyte interface, although the nature of this interface will, in general, dictate the magnitude of the measured overpotential.

When studying an anodic or cathodic working electrode in electrochemistry, one

would like to separate the η_W and η_C terms in order to know η_W for any fixed value of the current. This is done with the use of a reference electrode. It is worth reminding that a galvanostat always fixes the current between the working (W) and counter (C) electrodes at a desired value, whereas a potentiostat always fixes the current between the working and counterelectrodes so that the potential between the working (W) and reference (R) electrodes is at a fixed desired value U_{WR}. In either case, the current, I, is flowing between the working and counterelectrodes. In principle, no current passes through the reference electrode. In practice, usually a very small current passes through the reference electrode (this problem is more pronounced in solid-state electrochemistry) and an appropriate correction has to be made as follows.

One rewrites Eqs (9) and (10) but now considering the working (W) and reference (R) electrodes:

$$\eta_{WR} = U_{WR} - U^{o}_{WR} \qquad (11)$$

where η_{WR} is the working-reference overpotential and U^{o}_{WR} is the open-circuit ($I = 0$) value of the potential difference U_{WR}. Again η_{WR} consists of three terms:

$$\eta_{WR} = \eta_W + \eta_R + \eta_{ohmic,WR} \qquad (12)$$

Ideally no current flows through the reference electrode; therefore, $\eta_R = 0$ and $\eta_{ohmic,WR} = 0$ should be the case. In practice, the first assumption is usually good for reasonably nonpolarizable reference electrodes, since the parasitic uncompensated current flowing through the reference electrode is usually very small. The ohmic drop, however, between the working and reference electrodes, that is, $\eta_{ohmic,WR}$, may in general, and particularly in solid-state electrochemistry, not be negligible and must be determined using the current interruption technique, preferably in conjunction with a recording oscilloscope [7, 23]. The ohmic component decays to zero within less than 1 μs and the remaining part of η_{WR} is η_W (Fig. 14). In aqueous electrochemistry, the ohmic overpotential $\eta_{ohmic,WR}$ is minimized by using a Luggin capillary which brings the reference electrode as close as possible to the working electrode and not in the current path between the working and counterelectrodes in the electrolyte. Similarly, in solid-state electrochemistry, the reference electrode must be placed as close to the working electrode as possible to minimize $\eta_{ohmic,WR}$. This, however, is practically not easy when using flat electrolyte components (e.g., plates or disks) in which case the parasitic $\eta_{ohmic,WR}$ ohmic term must always be measured accurately [23]. Figure 15 shows typical current–overpotential curves (Tafel plots) for Pt (Fig. 15a) and Ag electrodes interfaced with YSZ and also acting simultaneously as heterogeneous catalysts for C_2H_4 oxidation. Figure 15(a) shows the importance of metal–gas–electrolyte three-phase-boundary (tpb) length, as the tpb length of the Pt film labeled R1 is 10^4 times larger than that of the film R2. The tpb length can be controlled by sintering during electrocatalyst preparation [23].

2.3.3 Cyclic Voltammetry

Cyclic voltammetry is still the most commonly used technique for the investigation of electrocatalysis. It involves the repeated imposition of a triangular potential function, against time, usually created by a function generator and applied via a potentiostat between the working and reference electrodes, while recording the

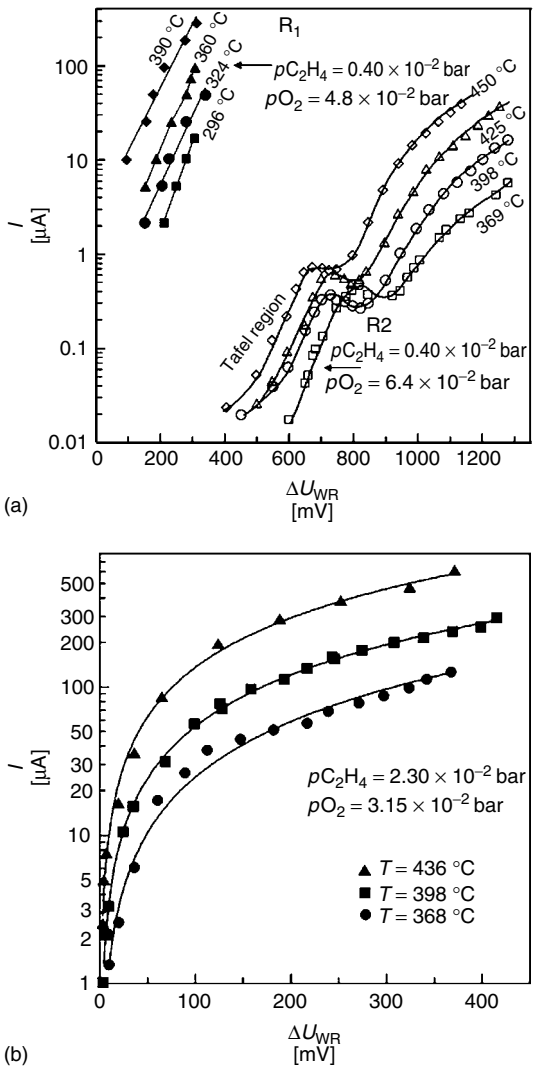


Fig. 15 (a) Typical Tafel plots for Pt catalyst–YSZ interfaces during C_2H_4 oxidation on Pt; the large difference in I_0 values between the two Pt films (labeled R1 and R2) is due to the higher calcination temperature of Pt film R2 versus Pt film R1 and concomitant pronounced decrease in the three-phase-boundary (tpb) length. (b) Effect of temperature on the Tafel plots and corresponding I_0 values of a Ag catalyst–YSZ interface during C_2H_4 oxidation on Ag. (Reprinted with permission from Ref. 23, Copyright 2003 by Kluwer/Plenum Publisher).

current generated between the working and counterelectrodes as a function of the applied potential. The theory of cyclic voltammetry is presented lucidly by Bard and Faulkner [7]. Typical cyclic voltammograms (CVs) of Pt in alkaline (1 M KOH) and of Pt and Au in acidic (0.5 M H_2SO_4) solutions [8, 38] are shown in Fig. 16.

The advantages of cyclic voltammetry versus stationary current–potential plots is twofold. First, all the difficulties related to the elimination of diffusion overpotential in stationary measurements are solved without having to use rotating disk or ring electrodes. Second, the problem of impurities accumulating on the electrode surface, which can severely affect the performance of the electrocatalyst during prolonged stationary measurements, is also, to a large extent, eliminated owing

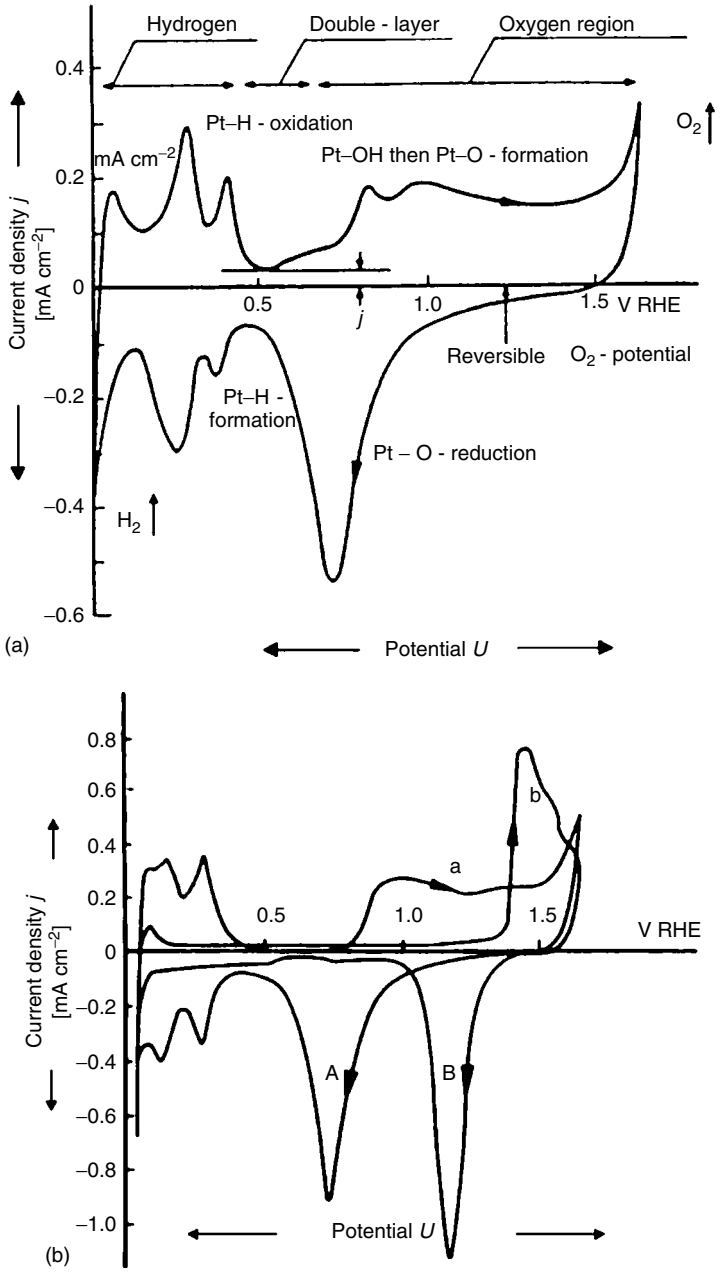


Fig. 16 (a) Cyclic voltammogram of polycrystalline platinum in 1 M KOH. The solution was purged with N_2, and the measurement was carried out at 20 °C with $|\nu| = 100$ mV s^{-1}. (b) Cyclic voltammogram in 0.5 H_2SO_4 at room temperature, $|\nu| = 100$ mV s^{-1} with nitrogen purging: (A) platinum and (B) gold. (Reprinted with permission from Ref. 8, Copyright 1998 by Wiley-VCH).

to electrochemical oxidation or reduction of these impurities. The disadvantage is, of course, that the accurate determination of the three parameters, i_0, a_a, and a_c, of the Butler–Volmer equation for a given electrocatalytic reaction, from a CV is, realistically, not feasible. One, instead, gets a qualitative and, in some cases, semiquantitative assessment of the activity of the electrocatalyst for a specific charge-transfer reaction from the deviation of the observed potential (at the peak maximum) from the thermodynamic equilibrium potential for the charge-transfer reaction under consideration. The smaller this deviation, the better the electrocatalyst.

Since these potential deviations (or overpotentials) from the equilibrium potential are usually equal and opposite, it follows that the smaller the potential difference between the anodic and cathodic peak, the better the electrocatalyst. If the two peaks are exactly symmetric, that is, opposite to each other, it follows that the charge-transfer reaction is significantly faster than the potential scan rate υ (mV s^{-1}).

The scan rate, υ, and the negative (cathodic) and positive (anodic) turnaround potentials, U_t^c and U_t^a, are the operating parameters of cyclic voltammetry. The former, υ, is usually chosen in the range 10 to 100 mV s^{-1}, whereas the latter in aqueous electrochemistry are chosen between the potentials for the hydrogen (reversible hydrogen electrode (RHE = 0) and oxygen (RHE = 1.23 V) evolution potentials. Similarly, in solid-state electrochemistry, U_t^c and U_t^a are chosen so that electrolysis of the solid electrolyte is avoided both at the working electrode (standard oxygen electrode (*SOE*) potential > -2.3 V in YSZ) and at the counterelectrode.

The basic features of a CV in aqueous solution, not containing an additional redox couple, are shown in Fig. 16. The potential scale (RHE) refers to a reversible H_2 electrode in the same solution. One distinguishes the double-layer capacitance region between 0.45 and 0.55 V, where the current is only used to charge the double layer [8]. The capacitance of the double layer, C_d, can be computed from $C_d = \upsilon / i$ where i has been normalized using the true electrode surface area. The value of C_d is of the order of $3 - 15$ μF cm^{-2}.

Above 0.55 V, the chemisorption of oxygen starts (oxide formation region) with the reaction:

$$Pt + OH^- \rightarrow Pt - OH + e^- \qquad (13)$$

followed by

$$Pt - OH + OH^- \rightarrow Pt - O + H_2O + e^- \qquad (14)$$

above ~0.8 V. Oxygen evolution, which is thermodynamically possible above 1.23 V, starts usually above 1.6 V, that is, there is a significant activation overpotential for O_2 evolution, and thus also for oxide reduction, as can be seen in Fig. 16, even when using Pt electrodes [8].

In the cathodic region ($I < 0$), after the oxide reduction peak

$$Pt - O + H_2O + 2e^- \rightarrow Pt + 2OH^- \qquad (15)$$

one obtains H chemisorption via

$$Pt + H_2O + e^- \rightarrow Pt - H + OH^- \qquad (16)$$

close to the thermodynamic potential (RHE = 0) of the H_2/H_2O couple, followed by H_2 evolution. At the beginning of the anodic scan ($I > 0$), one observes three Pt–H oxidation peaks, assigned to the reverse of reaction (16) and corresponding to three different H adsorption sites.

These hydrogen reduction/oxidation peaks near the low-potential end of the CV are used routinely for computing the

true surface area of polycrystalline electrodes by comparing their area with that obtained with Pt single-crystal electrodes.

By comparing Figure 16(a) and (b), one observes that the electrolyte has a rather minor effect on the CV of Pt, provided it does not introduce a redox couple. On the other hand, the CV of Au (Figure 16(b)) is substantially different: The H adsorption peaks are very severely suppressed, the double-layer region is very significantly expanded, and the oxide formation and reduction occur at much more positive potentials, indicative of the much weaker adsorption of hydrogen and oxygen on Au than on Pt.

When oxidation processes take place on the electrode surface, the CV becomes more complex [8] (Fig. 17). The oxidation of an organic molecule gives rise to a strong anodic ($I > 0$) peak and the anodic cleanup of the surface from the oxide,

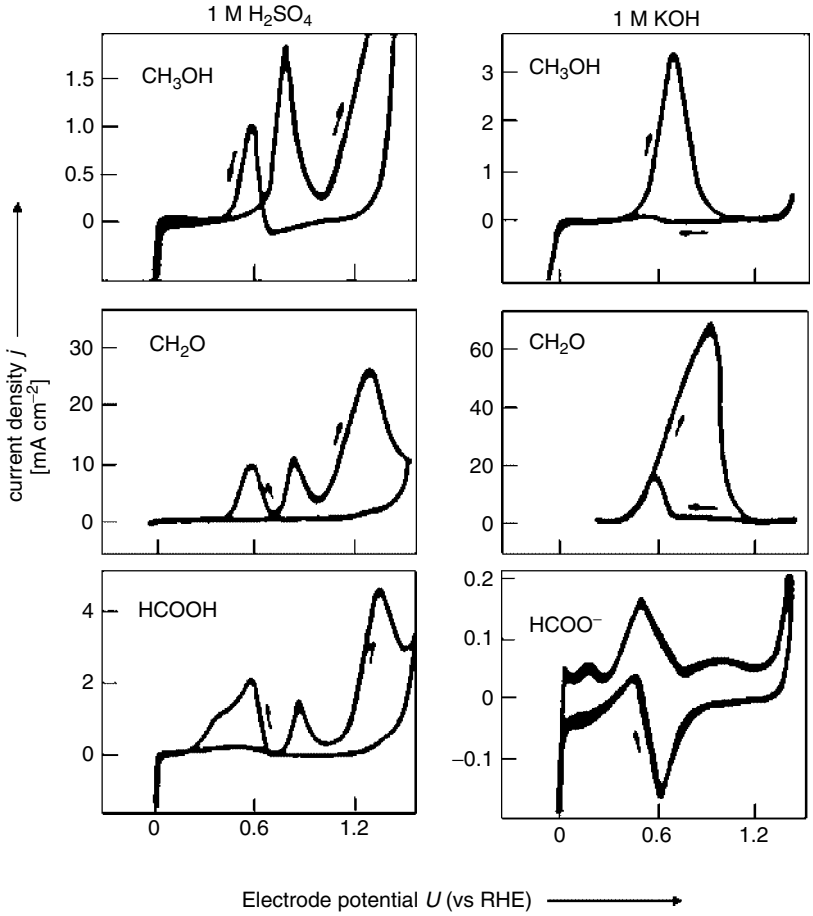


Fig. 17 Cyclic voltammogram of 1 M methanol, 1 M formaldehyde, and 1 M formic acid at smooth platinum electrodes in 0.5 H_2SO_4, 1 M KOH, 20 °C, all solutions purged with nitrogen to remove oxygen. (Reprinted with permission from Ref. 8, Copyright 1998 by Wiley-VCH).

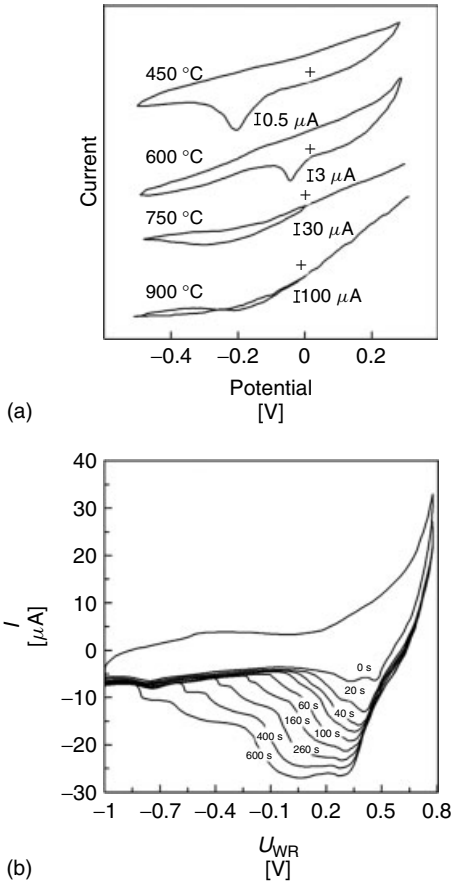


Fig. 18 (a) Cyclic voltammograms of Pt/YSZ electrodes in air at different temperatures at 50 mV s^{-1}. (Reprinted with permission from Ref. 39. Copyright © 1991 Elsevier) (b) Cyclic voltammograms of a Pt/YSZ film at 673 K and various holding times under UHV. Holding potential: 0.8 V; scan rate: 50 mV s^{-1}. (Reprinted with permission from Ref. 40, Copyright 1995 by the American Chemical Society).

which can passivate the surface, thus causing the disappearance of the oxygen reduction peak and the appearance of a second oxidation peak at more negative potentials in the reverse scan. The better electrocatalyst is obviously the one where the anodic oxidation peak appears as left as possible, that is, as close to the thermodynamic potential of the oxidation reaction as possible.

Solid electrolyte CVs bear several similarities and differences with aqueous CVs. In the examples in Fig. 18, the Pt/YSZ (an O^{2-} conductor) electrode is exposed to O_2, so there are obviously no hydrogen reduction/oxidation peaks and, quite often, no anodic oxidation peaks [23, 39, 40]. This is because the oxide (chemisorbed oxygen layer) is formed spontaneously from gaseous O_2 during the scan in a net catalytic (no charge transfer) step, thus not giving rise to an anodic oxidation peak. As a result, the only observable peak is the chemisorbed oxygen (oxide) reduction peak. Figure 18(b) shows that when the holding time at the anodic potential is increased, different types of surface oxides grow. This is quite important for studying the effect of EP or NEMCA effect. The first reduction peak corresponds to normally chemisorbed oxygen (forming the well-known $(2 \times 2)O$ overlayer on Pt(111)), whereas the second reduction peak corresponds to the reduction of more ionically bonded ''spillover'' oxygen originating from the YSZ support (rather than from the gas phase) and acting as a ''sacrificial promoter'' in the EPOC oxidations [23]. These observations have been confirmed using X-ray photoelectron spectroscopy (XPS), [23] and also $^{18}O_2$ [41] investigations.

2.3.4 AC-impedance Spectroscopy

The technique of AC-impedance spectroscopy is one of the most commonly used techniques in electrochemistry, both aqueous and solid [7, 8]. An AC voltage of small amplitude and frequency

f is applied between the working and reference electrodes, superimposed to the catalyst potential U_{WR}, and both the real (Z_{Re}) and imaginary (Z_{Im}) part of the impedance $Z(=dU_{WR}/dI = Z_{Re} + iZ_{Im})$ are obtained as a function of f (Bode plot, Fig. 19a). Upon crossplotting Z_{Im} against Z_{Re}, a Nyquist plot is obtained (Fig. 19b). One can also obtain Nyquist plots for various imposed U_{WR} values as shown in Fig. 20.

Owing to the small amplitude of the superimposed voltage or current, the current–voltage relationship is linear and, thus, even charge-transfer reactions, which normally give rise to an exponential current–potential dependence, appear as resistances, usually coupled with a capacitance. Thus, any real ohmic resistance associated with the electrode will appear as a single point, whereas a charge-transfer reaction (e.g., taking place at the tpb) will appear ideally as a semicircle, that is, a combination of a resistor and capacitor connected in parallel (Fig. 19).

Figures 19(a) and (b) show the Bode and Nyquist plots for a resistor, R_0, connected in series with a resistor, R_1, and a capacitor, C_1, connected in parallel. This is the simplest model which can be used for a metal–electrolyte or a metal–solid electrolyte interface. Note in Fig. 19(b) how the first intersect of the semicircle with the real axis gives R_0 and how the second intersect gives $R_0 + R_1$. Also note how the capacitance, C_1, can be computed from the frequency value, f_m, at the top of the semicircle (summit frequency), from $C_1 = 1/2\pi f_m R_1$.

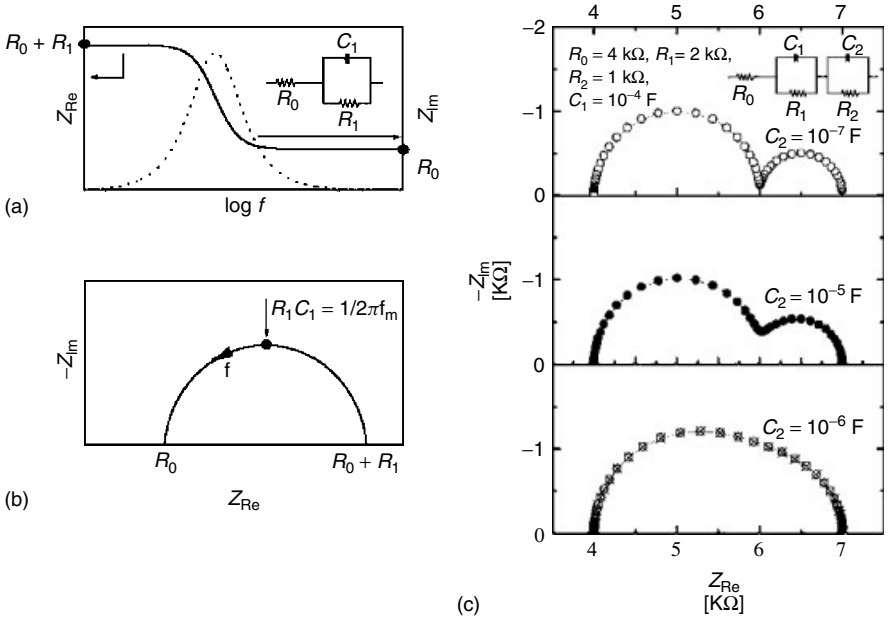


Fig. 19 Bode (a) and corresponding Nyquist plot (b) of the circuit, shown in inset, which is frequently used to model a metal/solid electrolyte interface. Effect (c) of capacitance C_2 on the Nyquist plot at fixed R_0, R_1, and R_2. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publisher).

Figure 19(c) shows Nyquist plots corresponding to a more complex resistor–capacitor network which is frequently found to describe the behavior of metal–solid electrolyte interfaces under electrochemical promotion conditions well. The left, high-frequency semicircle simulates the charge-transfer reaction at the tpb and the right, low-frequency semicircle simulates the charge-transfer reaction over the entire gas-exposed electrode surface. Figure 19(c) shows that the second semicircle appears clearly only when C_2 exceeds C_1 by 2 orders of magnitude.

Figure 20(a) exemplifies such a behavior of a Pd catalyst electrode deposited on YSZ and exposed to CH_4/O_2 mixtures [42]. The resistance R_0 is associated with the ohmic resistance of the electrode, whereas the semicircles labeled C_1 and C_2 are associated with the charge-transfer reaction

$$O^{2-} \rightleftharpoons O(a) + 2e^- \qquad (17)$$

taking place at the tpbs, YSZ–Pd–gas, (semicircle C_1), and over the entire Pd–gas interface (semicircle C_2). They correspond to capacitance values $C_{d,1}$ and $C_{d,2}$ computed from $C_{d,i} = 1/2\pi f_{m,i} R_i$ where $i = (1, 2)$ and $f_{m,i}$ is the frequency at the peak of semicircle i. The capacitance values are $C_{d,1} \approx 0.1 - 10\ \mu F\ cm^{-2}$ and $C_{d,2} \approx 500\ \mu F\ cm^{-2}$ (Fig. 20b).

The semicircle labeled C_2 is the ''backspillover oxygen semicircle''. It appears only at positive imposed U_{WR} values, that is, when O^{2-} is supplied to the catalyst surface. It corresponds to a capacitance $C_{d,2}$ again computed from $C_{d,2} = 1/2\pi f_{m,2} R_2$ and gives a $C_{d,2}$ value of $200\ \mu F\ cm^{-2}$. It is due to the charge-transfer reaction (17) now taking place over the entire gas-exposed electrode surface area. The dependency of $C_{d,1}$, and $C_{d,2}$ on potential is shown in Fig. 20(b).

For the experiments shown in Fig. 20(a) and (b), the ratio $C_{d,2}/C_{d,1}$ is on the average 2500, very close to the ratio N_G/N_{tpb} ($\approx$3570) where N_G is the gas-exposed electrode surface area and N_{tpb} is the ''surface area'' of the tpbs [23, 42]. These quantities were measured by surface titration and scanning electron microscopy (SEM), and other techniques described in detail in Refs 23, 41. Therefore, once N_G has been measured, AC-impedance spectroscopy allows for an estimation of the tpb length via

$$\frac{N_{tpb}}{N_G} = \frac{C_{d,1}}{C_{d,2}} \qquad (18)$$

Similar conclusions were derived from the work of Kek, Pejovnik, and Mogensen [43] who were the first to use AC-impedance spectroscopy for the detailed investigation of Pt, Au, and Ni films deposited in YSZ and exposed to mildly reducing environments (Fig. 20c–e). The low-frequency capacitance of the metal/YSZ interface is up to $300\ \mu F\ cm^{-2}$, manifesting the presence of a double layer at the metal/gas interface even under mildly reducing conditions ($\sim$1% H_2) (Fig. 20c,d). This double layer is stabilized even on Au by prolonged anodic polarization (Fig. 20e).

Thus, in the metal/YSZ systems of solid-state electrochemistry, AC-impedance spectroscopy provides concrete evidence for the formation of an effective electrochemical double layer over the entire gas-exposed electrode surface. The capacitance of this metal/gas double layer is of the order of $100–500\ \mu F\ cm^{-2}$ of superficial electrode surface area and of the order $2–10\ \mu F\ cm^{-2}$ when the electrode roughness is taken into account and, thus, the true metal/gas interface surface area is used, comparable to that corresponding to the metal/solid electrolyte double layer. Furthermore AC-impedance spectroscopy

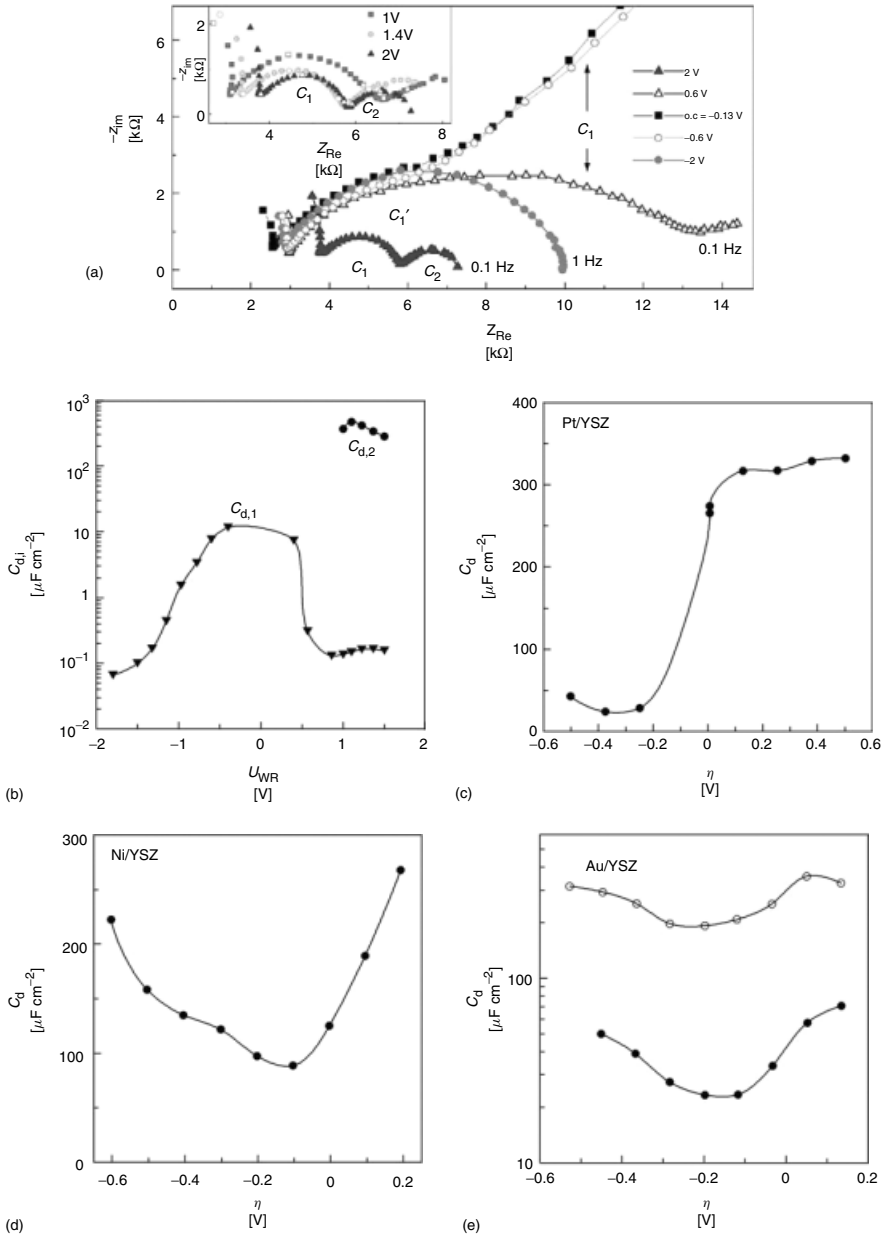


Fig. 20 (a) Complex impedance spectra (Nyquist plots) of the CH_4, O_2, and Pd|YSZ system at different Pd catalyst potentials. Open-circuit potential $U^{\circ}_{WR} = -0.13$ V. (b) Dependence on catalyst potential of $C_{d,1}$ and $C_{d,2}$. (Reprinted with permission from Ref. 23, Copyright 2003 by Kluwer/Plenum Publisher) (c) Double-layer capacitance as a function of overpotential of the systems Pt/YSZ, (d) Ni/YSZ, and (e) Au/YSZ, the latter before (●) and after (○) prolonged anodic overpotential application [42, 43]. (Reprinted with permission from Ref. 23, Copyright 2003 by Kluwer/Plenum Publisher).

permits estimation of the tpb length via Eq. (18) once the gas-exposed electrode surface area N_G is known.

2.3.5 Normalization of Active Surface

For the quantitative comparison of different electrocatalysts, it is necessary to know their true active surface area. The experimental techniques which allow for the determination of the true surface area of an electrocatalyst can be divided into ex situ and in situ techniques.

In the first group belong the techniques which are also used in heterogeneous catalysis for determining the surface area of catalysts. Two such techniques are widely used: The Brunauer–Emett–Teller (BET) method, based on the physical adsorption of N_2 or Ar at very low temperatures [8, 44] and the H_2 or CO chemisorption method [8, 44]. The first method leads to the total catalyst surface area, whereas the second leads to the specific (active metal) surface area. In the case of supported electrocatalysts (e.g., Pt/C electrocatalysts used as anodes in PEM fuel cells) the two techniques are complementary, as the former can lead to the total electrocatalyst surface area and the latter to supported metal (e.g., Pt) surface area.

In situ techniques are based primarily on the determination of the charge stored in the H adsorption region of the CV of the electrocatalyst (Fig. 21). As shown in this figure, the double-layer charge must be subtracted for a reliable measurement. The currently recommended normalization value, based on a 1:1 H to Pt chemisorption is 210 μC cm^{-2}, which roughly corresponds to the surface density of Pt(111) [38].

The use of cyclic voltammetry in conjunction with CO adsorption and CO stripping for estimating electrocatalyst surface area and CO coverage is shown in Fig. 22 for Pt and Rh electrocatalysts in acidic solution [45, 46].

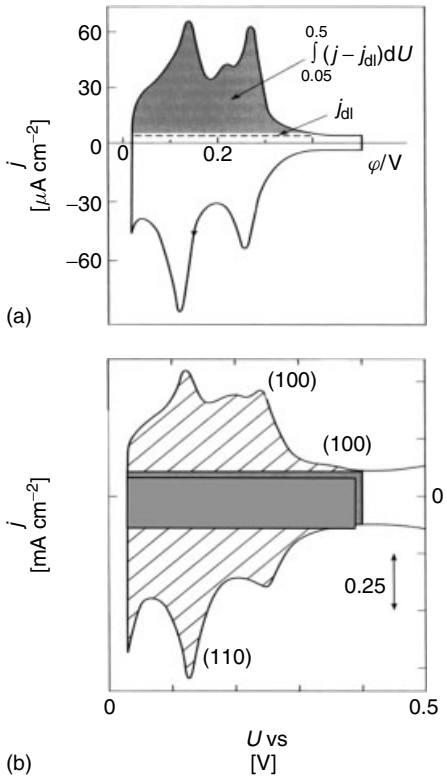


Fig. 21 (a) Cyclic voltammogram of a smooth platinum electrode in 0.05 M H_2SO_4 solution showing the principle of the method to integrate the electric charge for the hydrogen desorption. Sweep rate: 0.05 V s^{-1}. (b) Cyclic voltammogram of Pt/Vulcan dispersed electrode with a loading of 28 μg Pt cm^{-2}. The gray rectangle is the estimated double-layer charge and the Vulcan capacitance used to correct the charge due to the hydrogen adsorption–desorption. The solution was 0.5 M H_2SO_4 at 25 °C and the speed rate is 0.02 V s^{-1}. (Reprinted with permission from Ref. 9, Copyright 2003 by John Wiley & Sons Ltd).

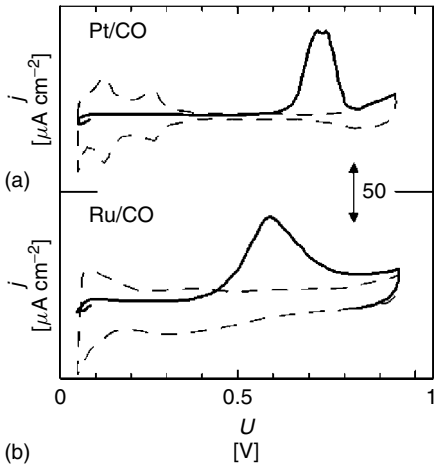


Fig. 22 Cyclic voltammogram for CO stripping at sputter-cleaned electrodes in 0.5 M H_2SO_4, dashed lines after CO removal. (a) Pt and (b) Ru. Immersion and adsorption at 250 mV, sweep rate: 20 mV s^{-1}. (Reprinted with permission from Refs 45, 46, Copyright 1994 by the American Chemical Society).

2.4 Theory of Electrocatalysis

2.4.1 Overpotentials and the Butler–Volmer Equation

The concept and theory of overpotential is of central importance in understanding electrocatalysis [47–55].

Operationally, the measured overpotential η of a working electrode (W), is the deviation of its potential from its open-circuit ($I = 0$) value and can be considered to be the sum of three terms, that is:

$$\eta_W = \eta_{ac,W} + \eta_{conc,W} + \eta_{ohmic,W} \quad (19)$$

where $\eta_{ac,W}$, $\eta_{conc,W}$, and $\eta_{ohmic,W}$ are the activation overpotential, the concentration (diffusion or mass transport) overpotential, and the ohmic overpotential respectively.

Once the ohmic overpotential, $\eta_{ohmic,W}$, has been eliminated or computed via current interruption, one is left with the activation and concentration (or diffusion) overpotentials only. The activation overpotential, $\eta_{ac,W}$, is due to slow charge-transfer reactions at the electrode–electrolyte interface and is related to current via the Butler–Volmer equation discussed below. A slow chemical reaction (e.g., adsorption, desorption, spillover [23]) preceding or following the charge-transfer step can also contribute to the development of activation overpotential.

The concentration overpotential $\eta_{conc,W}$ is due to slow mass transfer of reactants and/or products involved in the charge-transfer reaction. There exist simple equations for computing its magnitude in terms of mass-transfer coefficients or, more frequently, in terms of the limiting current I_l, which is the maximum current obtained when the charge-transfer reaction is completely mass transfer controlled. In aqueous electrochemistry, the concentration (or diffusion) overpotential can be minimized by the use of the rotating disk or rotating ring electrodes which, when rotated sufficiently fast, lead to very high mass-transfer coefficients and limiting currents I_l. Contrary to aqueous electrochemistry, where concentration overpotential is frequently important due to low reactant and/or product diffusivities in the aqueous phase, in solid electrolyte cells mass transfer in the gas phase is fast and, consequently, gaseous concentration overpotential is usually negligible, particularly in electrochemical promotion applications where the currents involved are usually very small.

The dependence of the activation overpotential on current in generally described

by the Butler–Volmer equation:

$$I/I_0 = i/i_0 = \exp(a_a F\eta/RT) - \exp(-a_c F\eta/RT) \quad (20)$$

where I_0 and i_o are the exchange current or exchange current density, respectively, and a_a and a_c are the anodic ($I > 0$) and cathodic ($I < 0$) transfer coefficients.

The usual procedure for extracting the exchange current I_0 is to measure η_W as a function of I and to plot $\ln I$ versus η_W (Tafel plot). Such plots were shown in Fig. 1 for aqueous electrolytes and in Fig. 15 for Pt and Ag catalyst electrodes interfaced with YSZ in solid-state electrochemistry.

When $|\eta| > 100$ mV, Butler–Volmer equation reduces to its ''high field approximation'' form, that is,

$$\ln(I/I_0) = a_a F\eta/RT \quad (21)$$

for anodic ($I > 0$, $\eta > 0$) operation and to

$$\ln(-I/I_0) = -a_c F\eta/RT \quad (22)$$

for cathodic ($I < 0, \eta < 0$) operation which lead to the well-known Tafel plots (Figs 1 and 15). Thus, by extrapolating the linear part of the $\ln|I|$ versus η plot to $\eta = 0$ one obtains I_0. The slopes of the linear parts of the plot give the transfer coefficients α_a and α_c. One can then plot I versus η and use the ''low field'' approximation of the Butler–Volmer equation which is valid for $|\eta| < 10$ mV, that is,

$$I/I_0 = (a_a + a_c)F\eta/RT \quad (23)$$

in order to check the accuracy of the extracted I_0, a_a, and a_c values.

The parameters a_a and a_c contain useful information about the reaction mechanism and their analytical computation using the symmetry factor, β (~0.5); the reaction mechanism is given in detail by Bockris [5].

Thus, for a multistep reaction

$$A + ne^- \longrightarrow B \quad (24)$$

and denoting the number of one-electron steps preceding and following the rate-determining step (rds) by $\vec{\gamma}$ and $\overleftarrow{\gamma}$, the number of times the rds is repeated by ν, and the number of electrons exchanged in the rds by r, one obtains

$$\alpha_a = (n - \vec{\gamma})/\nu - \nu\beta;$$
$$\alpha_c = \vec{\gamma}/\nu + \nu\beta;\ \alpha_a + \alpha_c = n/\nu \quad (25)$$

In extension of the definition of reaction order in heterogeneous catalysis, the order of an electrocatalytic reaction, a_i, with respect to a reactant, j, is defined as

$$a_i = \left(\frac{\partial \log \vec{i}}{\partial \log c_j}\right)_{j\neq k,\Delta U} = \left(\frac{\partial \log i_0}{\partial \log c_j}\right)_{j\neq k} \quad (26)$$

which implies that the comparison is made at constant potential U, or overpotential η, so that a_i expresses the variation of the exchange current with varying concentration c_j.

Therefore, from Eq. (26) it follows that the exchange current i_0, which is the measure of the activity of an electrocatalyst, is not a simple temperature-dependent constant, but also depends on reactant concentration.

Thus, in solid-state electrochemistry, the exchange current I_0 is known to be strongly dependent both on temperature and on gaseous composition. It increases with temperature with an activation energy which is typically 35–45 kcal mol^{-1} for Pt and 20–25 kcal mol^{-1} for Ag films deposited on stabilized zirconia [23].

The I_0 dependence on gaseous composition is usually complex. It has therefore

been shown [23, 56] that I_0 goes through a maximum with increasing p_{O2} at any fixed temperature for Pt|$ZrO_2(Y_2O_3)$ catalyst films. These results can be described adequately on the basis of Langmuir-type adsorption of oxygen at the tpb, that is,

$$\theta_O = K_O p_{O_2}^{1/2}/(1 + K_O p_{O_2}^{1/2}) \qquad (27)$$

where θ_O is the oxygen coverage. It can be shown [23, 56] that

$$I_0 \sim [\theta_O(1 - \theta_O)]^{1/2} \qquad (28)$$

or, equivalently,

$$I_0 \sim K_O p_{O_2}^{1/4}/(1 + K_O p_{O_2}^{1/2}) \qquad (29)$$

which satisfactorily explains the observed maxima and the fact that I_0 is proportional to $p_{O_2}^{1/4}$ for low p_{O2} and to $p_{O_2}^{-1/4}$ for high p_{O2} [23, 56]. According to this model, the I_0 maxima correspond to $\theta_O = 1/2$. It has been found, however, that for low T and high p_{O2} the situation becomes more complicated owing to the formation of surface Pt oxide PtO_2 [23].

When other gases are present in the gas phase in addition to O_2, then I_0 can be affected in two different ways: first, because θ_O may be affected because of a catalytic reaction and/or because of competitive chemisorption. Second, because these gases may react with O^{2-} at the tpb. In general, it is difficult to determine experimentally which one of these two factors is more important.

The exchange current I_0 is also an important parameter for the quantitative description of the electrochemical promotion (NEMCA effect). It has been found both theoretically and experimentally [23] that the order of magnitude of the absolute value $|\Lambda|$ of the Faradaic efficiency Λ defined from

$$\Lambda = \Delta r/(I/2F) \qquad (30)$$

where Δr is the electrochemically induced change in the catalytic rate and I is the applied current, can be usually estimated for any catalytic reaction from

$$|\Lambda| \approx 2Fr_0/I_0 \qquad (31)$$

where r_0 is the regular, that is, open-circuit, catalytic rate. When using Eq. (31) to estimate the order of magnitude of the enhancement factor Λ expected for a given catalytic reaction, one must use the I_0 value measured in the presence of the reacting gas mixture. The fact that I_0 increases exponentially with temperature in conjunction with the fact that Λ is inversely proportional to I_0 explains why EPOC is limited usually to temperatures below 600 °C [23].

As already noted, the performance of a galvanic cell (for electricity production) or of an electrolytic cell (for the production of chemicals) is dictated by the magnitude of the overpotentials both at the anode and at the cathode, η_a and η_c, by the ohmic drop between the anode and the cathode and by the diffusional overpotentials, $\eta_{d,a}$ and $\eta_{d,c}$, at the anode and cathode respectively, resulting from slow mass transfer at the anode and cathode. The latter are described by

$$\eta_{d,a} = \frac{RT}{F}\ln(1 - i/i_{\ell,a});$$
$$\eta_{d,c} = \frac{RT}{F}\ln(1 - i/i_{\ell,c}) \qquad (32)$$

where $i_{\ell,a}$ and $i_{\ell,c}$ are the diffusion-limited anodic and cathodic current densities, respectively. The computation of $i_{\ell,a}$ and $i_{\ell,c}$ for different electrode geometries are described in other chapters of this encyclopedia. These diffusional limitations, resulting from slow mass transfer relative to the electrocatalytic reaction, present

a serious problem in the study of electrocatalysis. As already noted, in order to minimize their effect during electrocatalytic measurements, the rotational disk and rotational ring electrodes have been developed. Their theory and experimental aspects are presented in other chapters of this encyclopedia.

2.4.2 Volcano Plots

The Sabatier principle of catalysis also finds extensive application in the area of electrocatalysis: reactants should be moderately adsorbed on the catalyst/electrocatalyst surface. Very weak or very strong adsorption leads to low electrocatalytic activity. This has been demonstrated repeatedly in the literature by the use of volcano plots (Figs 23–25). In these plots, the electrocatalytic activity is plotted as a function of the adsorption energy of the key reactant or some other parameter related to it in a linear or near-linear fashion, such as the work function of the metal [5], or the metal−H bond strength when discussing the H_2 evolution reaction (Fig. 24) [54] or the enthalpy of the lower-to-higher oxide transition when examining the O_2 evolution reaction (Fig. 25) [55].

Considering a general reaction:

$$A^+ + e \longrightarrow B(a) \qquad (33)$$

where the adsorbed intermediate B(a) is in equilibrium with gaseous B, and expressing the rate of the forward and reverse charge-transfer reactions, r_F and r_r respectively, using the Butler–Volmer equations

$$r_f = I_f/F = k_f(1-\theta_B)\exp\frac{\alpha F\eta}{RT} \qquad (34)$$

$$r_r = -I_r/F = k_r\theta_B\exp\left(-\frac{\alpha F\eta}{RT}\right) \qquad (35)$$

one obtains upon multiplication and setting $I_f = -I_f = I_0$

$$(I_0/F)^2 = k_f k_r \theta_B(1-\theta_B) \qquad (36)$$

therefore

$$I_0 = (k_f k_r)^{1/2} F[\theta_B(1-\theta_B)]^{1/2} \qquad (37)$$

One thus observes that the exchange current I_0 which expresses the electrocatalytic activity is maximized when $\theta_B \approx 0.5$. If one uses a Langmuir isotherm to relate θ_B with p_B [56], then Eq. (37) gives

$$I_0 = (k_f k_r)^{1/2} F \frac{K_B^{1/2} p_B^{1/2}}{(1+K_B p_B)} \qquad (38)$$

which is maximum for $K_B p_B = 1$, that is, for moderate values of the adsorption equilibrium constant K_B. Figure 23 (Ref. 57), shows the expected dependence use of I_0

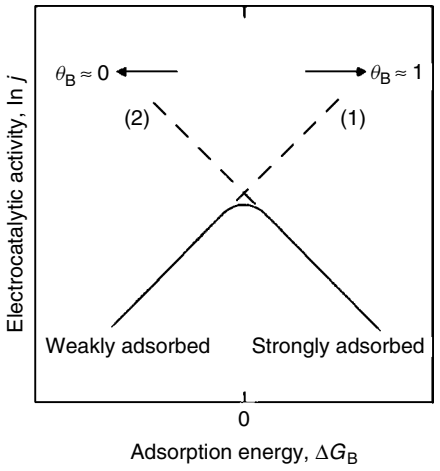


Fig. 23 Sketch of a "volcano" curve in electrocatalysis involving adsorption of B as an intermediate (Eq. 33) resulting from Eqs (38) and (41). (Reprinted with permission from Ref. 57, Copyright 2003 by John Wiley & Sons Ltd).

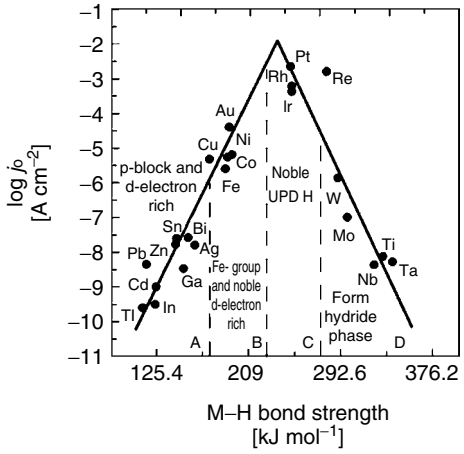


Fig. 24 Experimental volcano plot of the standard exchange current density, $\log j_0$, versus M−H bond energy on different metals. (Reprinted with permission from Ref. 55, Copyright 2000 by Elsevier Science).

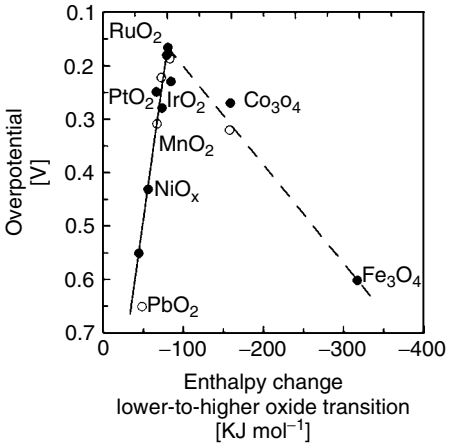


Fig. 25 Volcano plot of the overpotential for oxygen evolution versus the enthalpy of the lower-to-higher oxide transition (○) alkaline and (●) acidic solution. (Reprinted with permission from Ref. 54, Copyright 1984 by Elsevier).

on K_B and on the Gibbs energy of adsorption of the intermediate.

For example, in the case of dissociative O_2 adsorption in solid-state Pt/YSZ electrodes,

$$2O^{2-} \longrightarrow 2O(a) + 2e^-;$$
$$2O(a) \longrightarrow O_2(g) \qquad (39)$$

the same analysis gives

$$I_0 = (k_f k_r)^{1/2} F \frac{K_O^{1/2} p_{O_2}^{1/4}}{(1 + K_O p_{O_2}^{1/2})} \qquad (40)$$

which has been demonstrated to be in good agreement in solid-state electrochemistry for Pt/YSZ electrodes [23, 56].

From Eq. (40) and from the dependence of K_B on the energy of adsorption ΔG_B

$$\Delta G_B = -RT \ln K_B \qquad (41)$$

it follows that the dependence of the exchange current, I_0, on ΔG_A will be of the form given in Fig. 23. This type of plot is commonly termed *volcano plot* [57–65].

On the weak adsorption side of the volcano ($K_B p_B \ll 1$), Eq. (38) gives

$$\ln I_0 = \ln[(k_f k_r)^{1/2} F p_B^{1/2}] + \frac{\Delta G_B}{2RT} \qquad (42)$$

whereas on the strong adsorption side ($K_B p_B \gg 1$) it gives

$$\ln I_0 = \ln(k_f k_r)^{1/2} F / p_B^{1/2} - \frac{\Delta G_B}{2RT} \qquad (43)$$

that is, ideally the volcano plot is predicted to be symmetric with respect to ΔG_B (Fig. 23).

Since energies of adsorption are closely related to the chemisorption bond strength, one frequently uses the metal–adsorbate bond strength as the abscissa of volcano plots (Fig. 24).

This particular volcano plot is indeed nearly symmetric and demonstrates that Pt, Rh, Ir, and Re have the optimal M−H bond strength for maximum activity (exchange current density) for the hydrogen evolution/hydrogen oxidation reaction (HER/HOR).

Similarly, the overpotential for the O_2 evolution reaction is minimized (i.e., the electrocatalytic activity is maximized) for intermediate values of the enthalpy of the lower-to-higher oxide transition (Fig. 25).

2.4.3 Transition State Approaches

Starting from the pioneering work of Polanyi and Horiuti [3], the first theoretical approaches to electrocatalysis were based on the transition state theory. Thus, the rate of the reduction reaction

$$A^+ + e \longrightarrow A \tag{44}$$

can be expressed as

$$r_{red} = c_{A^+} c_e k_{red} \exp\left[\frac{-\Delta G^{\neq o}_{red}(\varphi_M)}{RT}\right] \tag{45}$$

where c_{A^+} is the concentration of H^+ at the OHP, c_e is the concentration of electrons in the metal electrode, k_{red} is a constant with dimensions of frequency, and $\Delta G^{\neq o}_{red}$ is the standard free energy of the transition state which is a function of the Galvani potential inside the metal, φ_M. This leads [5, 6, 66] to the single-electron exchange ($n = 1$) form of the Butler–Volmer equation:

$$i = i_0\left[\exp\left(\frac{\beta F\eta}{RT}\right) - \exp\left(\frac{(1-\beta)F\eta}{RT}\right)\right] \tag{46}$$

where $\eta(= \varphi_M - \varphi^o_M)$ is the overpotential and φ^o_M is the value of φ_M when $i = 0$. In the context of the transition state theory, the symmetry factor, β, represents the fraction of electronic charge remaining on the electrode at the transition state. The value of the symmetry factor β is usually near 0.5 and it also physically expresses the fraction of the potential difference across the double layer which diminishes the chemical barrier for an ion undergoing reduction (Fig. 26).

The transition state approach leads in a natural way to the Butler–Volmer equation, but is relatively weak in its predictive properties regarding the exchange current, i_0, which is proportional to the frequency factor k_{red} and to $\exp(-\Delta G^{\neq o}_{red})$. The latter is quite closely related to the enthalpy and standard entropy of formation of the adsorbed reduction product or intermediate, and this is one main reason for the very intense modern efforts to develop predictive theoretical tools for the ab initio computation of adsorption energies at

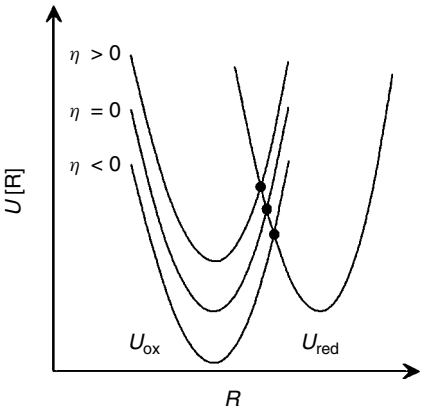


Fig. 26 Scheme of diabatic (nonadiabatic) potential energy surfaces $U(R)$ as functions of coordinate R for the process $Ox + e \rightarrow Red$ at different overpotentials η. The intersection of the curves corresponding to the oxidized U_{ox} and reduced U_{red} states are marked with dots. (Reprinted with permission from Ref. 34, Copyright 2003 by John Wiley & Sons Ltd).

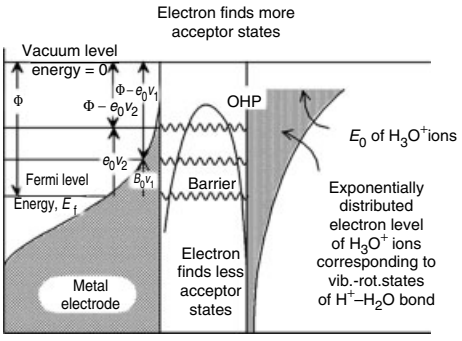
the metal–gas and metal–solution interfaces [53, 58].

2.4.4 Quantum Mechanics–based Models

Generally speaking, since charge transfer at the metal–solution interface is a dynamic process which involves electron tunneling [5, 17, 66], and for some systems such as H_2 oxidation also involves proton tunneling [6, 17], one would expect that a full treatment would involve the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial\Psi}{\partial t} = H\Psi \qquad (47)$$

where H is the Hamiltonian and Ψ is the wave function of the system. There are two common approaches to deal with this equation for electron transfer. One involves time-dependent perturbation theory [66, 67] and the other involves the adiabatic approximation [66, 68–75]. The electrochemical theory for electron transfer usually refers to the ground state and, for the case where no chemical bonds are formed or broken, utilizes the theory of outer-sphere reactions, developed by Marcus [68], applied specifically to the metal–solution interface, and termed *electrochemical outer-sphere reaction theory*.



In its simpler form, one considers a 1-D diabatic (intersecting) potential energy surface (PES) and expresses the reduction electron current density as

$$j_c = \alpha\int \rho(\varepsilon) f(\varepsilon) \times \exp\left[-\frac{(\lambda - \varepsilon + \varepsilon_o\eta)^2}{4\lambda k_b T}\right] d\varepsilon \qquad (48)$$

where α is a constant, λ is the energy of solvent reorganization, $\rho(\varepsilon)$ is the electronic density of states of the metal and $f(\varepsilon)$ is the Fermi–Dirac electron distribution function. The proportionality factor is proportional to the tunneling probability of the electron between the electrode and the oxidized species.

Similarly, for the anodic current one obtains

$$j_a = \alpha\int \rho(\varepsilon)(1 - f(\varepsilon)) \times \exp\left[-\frac{(\lambda + \varepsilon - \varepsilon_o\eta)^2}{4\lambda k_b T}\right] d\varepsilon \qquad (49)$$

The first visualization of the electron-tunneling process at the metal–solution interface is due to Gurney [17] (Fig. 27). The probability of the availability of end tunneling states at the OHP was considered to follow Boltzman's Law.

A more modern visualization of the factors entering the calculation of the anodic and cathodic currents has been developed by Gerischer and Schmickler (Fig. 28) [66, 69–71]. In this diagram, W_{ox} and W_{red} are proportional to the Gaussian terms in Eqs (48) and (49), corresponding

Fig. 27 Schematic diagram of the potential energy barrier at the metal–solution interface. The electron energy levels and their distribution in the electrode and in the solution are shown [17]. (Reprinted with permission from Ref. 17, Copyright 1979 by Plenum).

to the initial and final states of the reactant, whereas the $f(\varepsilon)$ term expresses the density of states in the metal electrode itself. At equilibrium ($\eta = 0$), $|j_a| = |j_c|$. The current density j_c is determined by the overlap of W_{ox} with the occupied states of the metal, whereas the current density j_a is determined by the overlap between W_{red} and the empty states in the metal. The application of a negative overpotential shifts both W_{ox} and W_{red} down relative to the Fermi level of the metal (Eqs 46 and 47), and this has two effects. First, it increases the overlap between W_{ox} and the occupied states, thus increasing the anodic current and, second, it decreases the overlap between W_{red} and the empty states, thus decreasing the cathodic current.

More recent work [72–81] has addressed the key theoretical aspects related to adsorption and electron-transfer process for important electrocatalytic system, such as, for example, CO adsorption and oxidation on transition metal [76] (Fig. 29).

An important practical development in practical ab initio quantum chemistry and electrochemistry in recent years has been the application of the density-functional theory (DFT) methods for the calculation of properties of large ensembles of atoms [82–84]. This is because the classical Hartree–Fock methods are extremely time consuming when large systems are involved. DFT calculations can compute binding energies of atoms and molecules with an accuracy 10–20 kJ mol^{-1}, which is a reasonable first approximation. In DFT calculations, the energy of the quantum chemical system is calculated directly from a single function, the electronic density. The functional transformation of electronic density to energy is not known, but good approximations can be obtained using the generalized gradient approximation (GGA) functionals [82–84]. The DFT/GGA methods are quite successful in treating chemical bonds, since they account for nonlocal electron correlation,

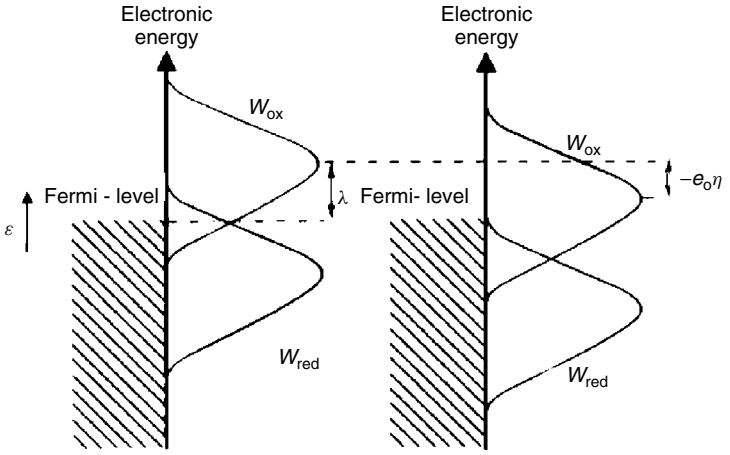


Fig. 28 Distribution functions W_{ox} and E_{red} at equilibrium and under application of a cathodic overpotential η. These distribution functions represent the contributions to the current by the oxidized/reduced states in the solution, originated by electron transfer from/to the energy state ε in the metal. (Reprinted with permission from Ref 22, Copyright 1996 by Oxford University Press).

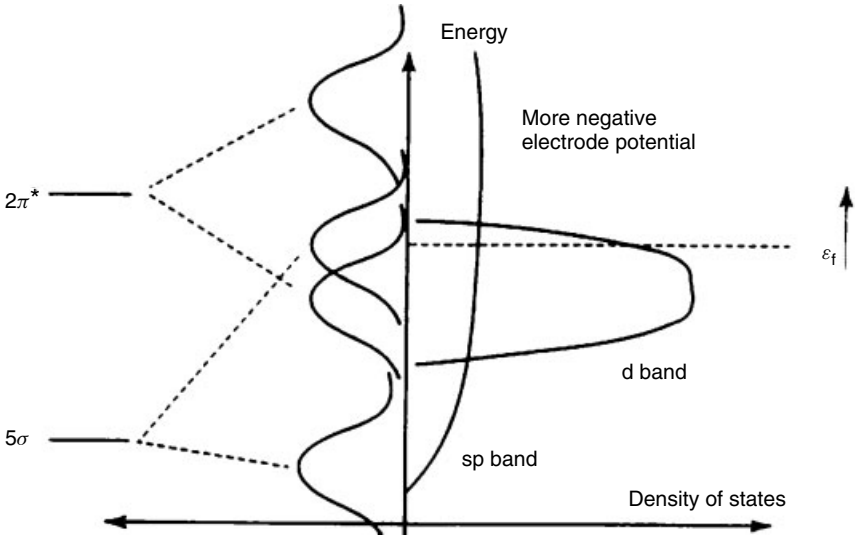


Fig. 29 Model of the orbital interaction diagram for CO adsorbed on transition metal surfaces. (Reprinted with permission from Ref. 76, Copyright 2003 by John Wiley & Sons Ltd).

which is more difficult to include in Hartree–Fock methods.

In treating molecule–surface interactions, two main approaches exist in the literature [26, 85–89]. The first considers the surface as a small (10–50) cluster of atoms [26, 87]. The second treats the surface as a slab in which a certain ensemble of atoms is repeated periodically [85–88].

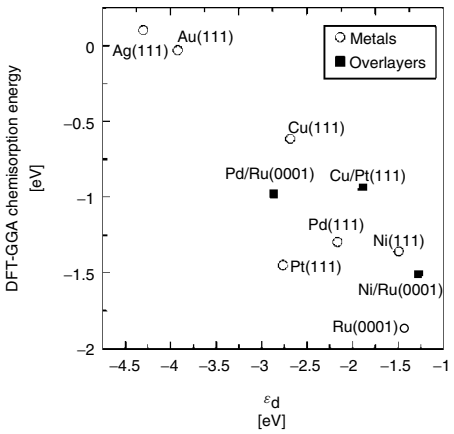


Fig. 30 Relationship between the center of the d band of the substrate relative to its Fermi level and the full DFT/GGA chemisorption energies for a number of metal systems. (Reprinted with permission from Ref. 87, Copyright 1996 by American Physical Society).

An important result recently obtained by Hammer and Norskov using extensive DFT/GGA calculation is that energies of adsorption of molecules, such as CO (Fig. 29), can be related in a linear fashion to the energy of the center of the d band, ε_d, relative to the Fermi level of the metal (Fig. 30) [87]. The closer the center of the d band relative to the Fermi level of the metal, the stronger the binding propensity with molecules, such as CO, where enhanced electron backdonation leads to stronger binding to the surface.

These considerations are very useful for rationalizing the observed electronic trends upon alloying Pt with Ru, Sn, or Mo in the search for CO-tolerant anodes for

PEM fuel cells. It is known that with several of these alloys the oxidation mechanism is bifunctional, for example

$$CO(Pt) + OH(Ru) \longrightarrow CO_2 + H^+ + e^- \quad (50)$$

but it is also well established that electronic effects are quite important [84–95], for example, that the binding of CO both on Pt and on Ru is altered significantly upon alloying [84].

2.4.5 Promotional Effects

As in heterogeneous catalysis [44], small coverages of electropositive or electronegative adatoms, acting as promoters or poisons, can have very significant effects in the performance of electrocatalysts, as they can significantly modify the metal work function [10, 23, 44] and, thus, its chemisorptive propensity [10, 23, 44].

The presence of such adatoms modifies the electric field in the double layer which strongly interacts with the adsorbed dipoles [23, 26, 27]. The effect of the electric field on the binding energy (BE) of adsorbates has been investigated by Pacchioni and Illas [26] and by Koper and van Santen [84, 92] using metal clusters. In general, a positive field weakens the BE of predominantly electron acceptor adsorbates, such as CO on clean Pt [23, 84] adsorbing on atop positions, but the trends can become more complex when examining adsorption in threefold hollow sites [84].

The presence of adatoms or skin layers of alloys causes significant changes in the work function of electrocatalysts and, therefore, in their chemisorptive and electrocatalytic properties. Watanabe and coworkers [96–98] have carried out systematic investigations on the effect of such skin layers on the XPS BE shifts of Pt 4d and 4f photoelectrons [98] (Fig. 31), and have correlated such shifts with the change in the coverage of adsorbed CO and change in electrocatalytic activity for the oxidation of H_2, in the presence of CO (Fig. 32) [96, 98]. This very important experimental work has shown conclusively (Figs 31 and 32) that CO-tolerant alloy electrocatalysts are characterized by a positive BE shift, whereas electrocatalysts that are not CO tolerant exhibit a negative BE shift.

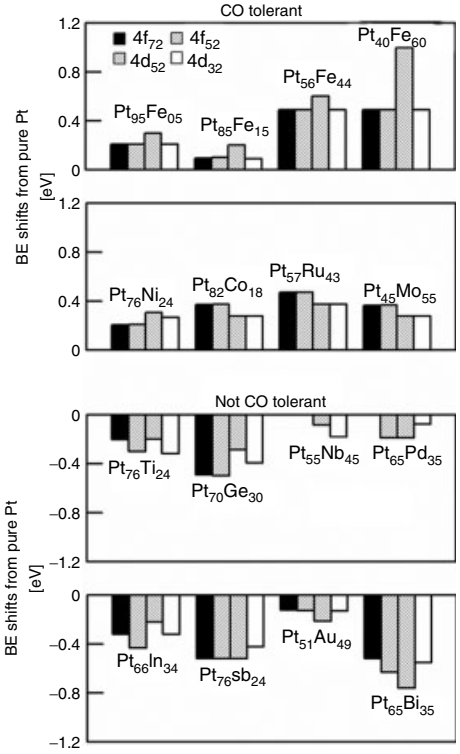


Fig. 31 Binding energy shifts at 4d and 4f in Pt skin layers of various CO-tolerant and non-CO-tolerant alloys [97, 98]. (Reprinted with permission from Refs. 97, 98, Copyright 2001 by The Royal Society of Chemistry on behalf of the PCCP Owner Societies).

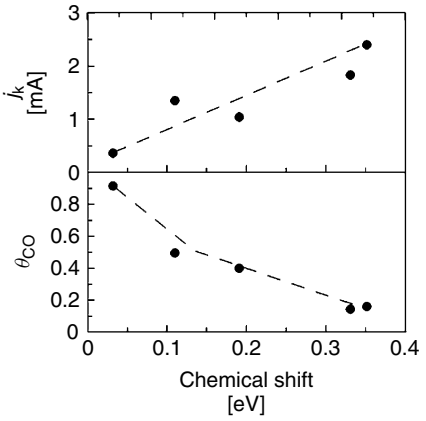


Fig. 32 Effect of chemical shift of Pt $4d_{5/2}$ binding energy in Pt skin layer, prepared on a bare $Pt_{50}Fe_{50}$ alloy surface by sputtering deposition, onto (top) the kinetically controlled H_2 oxidation current (J_k) and (bottom) the steady-state CO coverage (θ_{CO}) at room temperature in 0.1 M $HClO_4$ solution saturated with 100 ppm CO/H_2 balance [96, 98]. (Reprinted with permission from Ref. 96, Copyright 2003 by John Wiley & Sons Ltd).

A positive BE shift implies a more "negatively charged" Pt atom, and thus CO chemisorption is weaker (better electrocatalytic activity) on such alloy electrocatalysts. This is consistent with the predominantly electron acceptor nature of CO adsorption on most metals [23] and with the recently established rules of electrochemical promotion [23, 99] and of classical promotion in heterogeneous catalysis [100] which predict that CO binding is weakened on high work function, that is, "negatively charged" surfaces.

Another demonstration of these simple rules [23, 99, 100] is presented in Figs 33 and 34. Electronegative adatoms, such as S and Se, cause a dramatic increase in the rate of CO oxidation on Pt [101–104], similar to the case of CO-poisoned catalysts in heterogeneous catalysis [23, 44]. Both the EDL model of electrochemical promotion [23, 27] and rigorous ab initio quantum mechanical calculations [26] have shown that these pronounced promotional effects are due to the strong repulsive lateral interactions between adsorbed CO and coadsorbed electronegative adatoms [23, 26, 27].

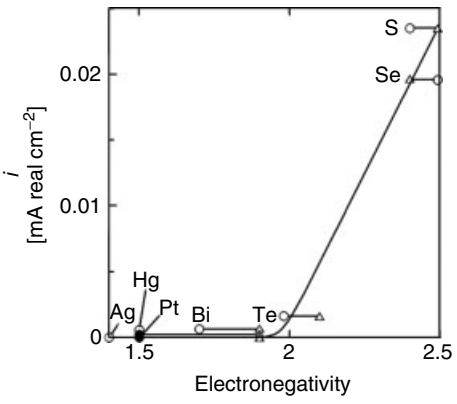


Fig. 33 Dependence of the maximum current density for CO oxidation at 0.45 V in 0.5 M H_2SO_4 on the electronegativity of the adatom. Electronegativity values from Pauling (Δ) and from Rochow (○) [102]. (Reprinted with permission from Ref. 102, Copyright 1985 by Elsevier Science).

2.5 The Hydrogen Oxidation/Evolution Reaction

The HOR and HER are the most thoroughly studied electrocatalytic systems [105–114]. Excellent reviews are given by Breiter [115], Lasia [116], and Markovic [117]. Conway and Bockris first established, in 1957, a linear relation between log i_0 and the metal work function Φ [118]. Parsons [60] and Gerischer [119]

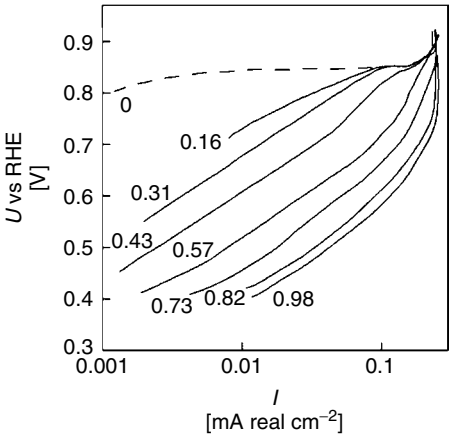


Fig. 34 Semilogarithmic plot of potential–current curves for CO oxidation on S-covered Pt electrodes in CO-saturated 0.5 M H_2SO_4 solution at 40 °C. Values of coverage θ_S^{Pt} for each curve are indicated. (Reprinted with permission from Ref. 102, Copyright 1985 by Elsevier Science).

independently used transition state theory to establish a linear relation between log i_0 and the chemical potential (Gibbs energy) of adsorbed hydrogen. The general form of this relation leads to volcano plots, as already shown in Fig. 24. Trasatti later showed that the same volcano plots apply both to metal and oxide electrocatalysts [120, 121].

These volcano plots show that Pt is the best electrocatalyst owing to its moderate strength of adsorption of hydrogen atoms. Adsorbed hydrogen atoms (Pt–H or frequently termed $Pt\text{–}H_{upd}$ where upd denotes underpotential deposition) is formed either from protons or from H_2O molecules:

$$\mathrm{Pt} + \mathrm{H_3O^+} + \mathrm{e^-} \rightleftarrows \mathrm{Pt\text{–}H_{upd}} + \mathrm{H_2O}\ (\mathrm{pH} < 7) \qquad (51)$$

$$\mathrm{Pt} + \mathrm{H_2O} + \mathrm{e^-} \rightleftarrows \mathrm{Pt\text{–}H_{upd}} + \mathrm{OH^-}\ (\mathrm{pH} > 7) \qquad (52)$$

The enthalpy of adsorption of H_{upd} on Pt(111) is not constant, but varies linearly from −42 to −22 kJ mol^{-1} as the coverage θ_H increases from 0 to 0.6 [117]. The HOR proceeds via the dissociative adsorption of H_2

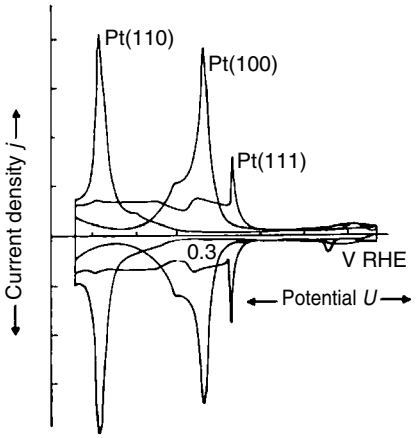


Fig. 35 Cyclic voltammogram of platinum single-crystal surfaces in the hydrogen adsorption region in 0.5 M H_2SO_4 under nitrogen purging; $|\nu| = 50$ mV s^{-1} [117, 122]. (Reprinted with permission from Ref. 122, Copyright 1991 by EDP Sciences).

$$\mathrm{H_2} + \mathrm{Pt} \rightleftarrows 2\,\mathrm{Pt\text{–}H} \qquad (53)$$

followed by two parallel routes in acidic solutions:

$$\mathrm{Pt\text{–}H} \longrightarrow \mathrm{Pt} + \mathrm{H^+} + e^- \qquad (54)$$

$$\mathrm{H_2} + \mathrm{Pt} \longrightarrow \mathrm{H^+} + \mathrm{Pt\text{–}H} + e^- \qquad (55)$$

where the relative importance of the two steps (Volmer–Tafel and Volmer–Heyrovsky) depends on the pH of the solution.

Similar to heterogeneous catalysis, the HER/HOR system is structure sensitive as shown by the CVs in Fig. 35 [122] and

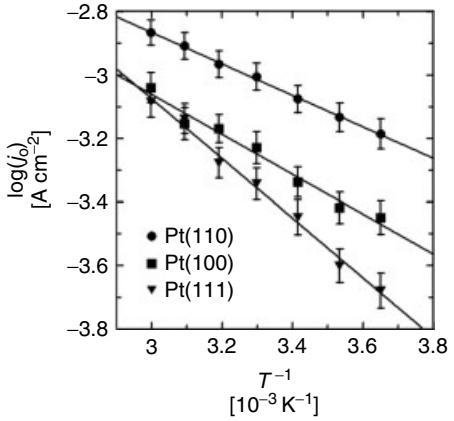


Fig. 36 Arrhenius plots of the exchange current densities (i_0) for the HER/HOR on Pt(hkl) in acid solution. (Reprinted with permission from Ref. 117, Copyright 2003 by John Wiley & Sons Ltd).

the steady-state log i_0 versus T plots in Fig. 36 [117], which demonstrate that the highest density Pt(111) surface is the less reactive one.

This structure sensitivity has been studied in detail by Markovic [117] and is also demonstrated in Fig. 37 which shows steady-state current–potential curves obtained in alkaline environment on different crystallographic Pt planes and also depicting the possible position of H_{upd} and H_{opd} on the different surfaces. It is interesting to note that the anodic current passes through a maximum near 0.4 V (versus RHE) which is primarily due to the very low coverage of H_{upd} at more positive potentials.

2.6 Oxygen Reduction/Evolution Reaction

This electrocatalytic system is of great importance [123–136] both for efficient fuel cell operation [123, 125] and for efficient electrolysis [123]. The progress in understanding oxygen electrocatalysis has been reviewed by Yeager [129], Appleby [130],

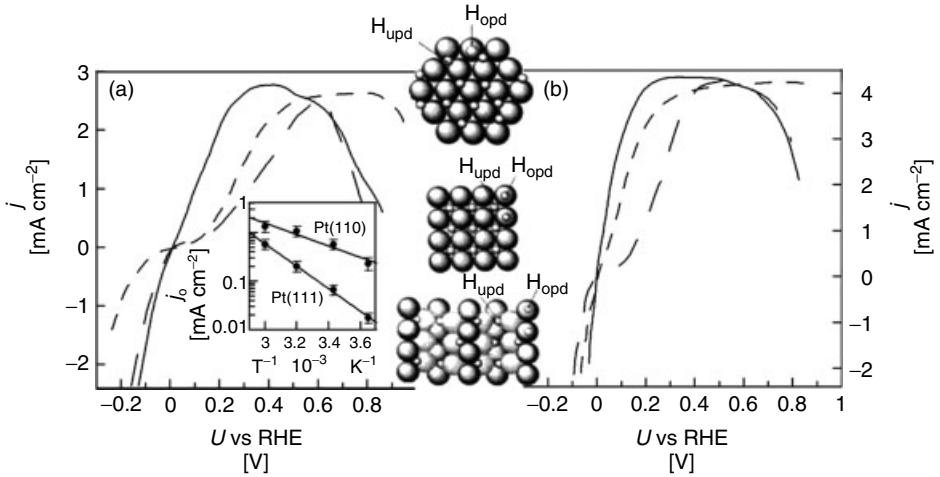


Fig. 37 (a) Polarization curves for the HER and the HOR on Pt(111) in 0.1 M KOH at 2500 rpm, 274 K, and 20 mV s^{-1}. (b) Polarization curves for the HER and the HOR on Pt(111) in 0.1 KOH at 2500 rpm, 333 K, and 20 mV s^{-1}. (Reprinted with permission from Ref. 117, Copyright 2003 by John Wiley & Sons Ltd).

Damjanovic [131], Kinoshita [132], Bockris [92], Adjic [134], MacDougall [123], and Ross [124] among others for aqueous systems and by McEvoy [28] and Ivers-Tiffée [125] for solid-state electrochemical systems.

The oxygen reduction reaction (ORR) at the cathode of an electrochemical cell involves the transfer of four electrons from the cathode to an O_2 molecule followed by removal of the reduced product from the interface. Reduction of an O_2 molecule to water or to hydroxyl anions at an electrode/aqueous electrolyte interface is to be distinguished from reduction of O_2 to $2O^{2-}$ at the surface of an oxide or at an electrode/solid electrolyte interface in solid-state electrochemistry. In the latter case, the ORR is generally different at the surface of a noble metal like Pt from that at the surface of a metallic oxide.

The ORR requires a weakening of the oxygen–oxygen bond of the O_2 molecule and therefore a strong interaction with the surface of the catalytic electrode. A noble metal like Pt has too large a work function to be able to donate two electrons to an adsorbed oxygen atom to create O^{2-} ions at the surface; but it can transfer one electron per oxygen atom, and this electron is charge compensated by a proton from the solution. If both oxygen atoms of the O_2 molecule are adsorbed on the surface, a bidentate adsorption occurs [137]. Thus, in acid solution

$$O_2 + 2e^- + 2H^+ \longrightarrow 2OH_{ads} \quad (56)$$

and in alkaline solution

$$O_2 + 2e^- + 2H_2O \longrightarrow 2OH_{ads} + 2OH^- \quad (57)$$

If the molecule is adsorbed end-on, only one electron is transferred. In *acid* solution

$$O_2 + e^- + H^+ \longrightarrow HO_{2,ads} \quad (58)$$

and in alkaline solution

$$O_2 + H_2O + e^- \longrightarrow HO_{2,ads} + OH^- \quad (59)$$

These two different types of adsorbed species lead to two different ORR pathways, one to a direct reduction to $2H_2O$ or $4OH^-$ and the other to a peroxide H_2O_2 or HO_2^- intermediate.

Thus, in acidic solutions, the first pathway proceeds via

$$2OH_{ads} + 2H^+ + 2e^- \rightarrow 2H_2O \quad (60)$$

and gives the overall direct reaction:

$$O_2 + 4H^+ + 4e^- \longrightarrow 2H_2O, \quad \Delta U^o = 1.23\ V \quad (61)$$

where the standard potential ΔU^o is referred to the standard hydrogen electrode (SHE). The second pathway proceeds via

$$HO_{2,ads} + e^- + H^+ \longrightarrow H_2O_2 \quad (62)$$

and leads to the overall reaction:

$$O_2 + 2H^+ + 2e^- \longrightarrow 2H_2O_2, \quad \Delta U^o = 0.682\ V \quad (63)$$

where

$$H_2O_2 + 2H^+ + 2e^- \longrightarrow 2H_2O, \quad \Delta U^o = 1.77\ V \quad (64)$$

may not be accessed [137]. Therefore, the indirect reaction leads to the generation of, usually undesirable, peroxide species and to a lowering of the open-circuit potential.

In alkaline solutions, the first pathway proceeds via

$$2OH_{ads} + 2e^- \longrightarrow 2OH^- \quad (65)$$

and gives the overall direct ORR:

$$O_2 + 2H_2O + 4e^- \longrightarrow 4OH^-, \quad \Delta U^o = 0.401\ V \tag{66}$$

whereas the second pathway:

$$HO_{2,ads} + e^- \longrightarrow HO_2^- \tag{67}$$

leads to the overall reaction:

$$O_2 + H_2O + 2e^- \longrightarrow HO_2^- + OH^-, \quad \Delta U^o = -0.076\ V \tag{68}$$

with

$$HO_2^- + H_2O + 2e^- \longrightarrow 3OH^-, \quad \Delta U^o = 0.88\ V \tag{69}$$

In solid-state electrochemical cells, for example those utilizing YSZ, the reduction of O_2 does not generally involve OH formation, but with noble metal electrodes proceeds via

$$O_2 + 2e^- \longrightarrow 2O_{ads}^- \tag{70}$$

$$2O_{ads}^- + 2e^- + 2\square_{el}\bullet \longrightarrow 2O_{el}^{2-} \tag{71}$$

where O_{ads}^- is adsorbed on the Pt from where it spills over to a vacancy, $\square_{el}$, in the solid electrolyte [137].

The reverse of reaction (71) is of particular importance in the area of EPOC [23], since the O_{ads}^- formed by anodic polarization of the Pt/YSZ interface is different from the O_{ads} formed by gaseous O_2 and acts as a sacrificial promoter for catalytic reactions taking place at the metal/gas interface where the O_{ads}^- species migrates (backspillover) from the solid electrolyte [23].

When a reducible metal oxide is used as the cathode, for example, perovskites such as $La_xSr_{1-x}MnO_3$ used in solid oxide fuel cells [28, 125], then, denoting by M the reducible metal, one can formulate the O_2 reduction pathway by

$$M^{(m-1)+} - \square - M^{(m-1)+} + O_2 \longrightarrow M^{m+} - (O_2)^{2-} - M^{m+} \tag{72}$$

followed by migration of the extra, active, oxygen over the surface to another vacancy where it also gets reduced as follows:

$$O_{ads} + M^{(m-1)+} - \square - M^{(m-1)+} \longrightarrow M^{m+} - O^{2-} - M^{m+} \tag{73}$$

followed by

$$2M^{m+} + 2e^- \longrightarrow 2M^{(m-1)+} \tag{74}$$

The O_2 evolution reaction which is of great importance in electrolytic cells, proceeds in general via the reverse of the above reactions. Owing to the principle of microscopic reversibility, it is obvious that active O_2 reduction electrocatalysts are equally active for the O_2 evolution reaction.

The use of mixed ionic-electronic conductors (MIECs) as ORR electrocatalysts is quite common in solid-state electrochemistry [125], because the reaction zone is extended over the entire electrode/gas interface, contrary to the case of metal electrodes where the reaction is, to a large extent, restricted to the tpb zone [23].

In general, finding efficient electrocatalysis for the oxygen reduction/evolution system is much more demanding than for the H_2 oxidation/evolution system. One reason is the much stronger chemisorption of O atoms on metals, in general, which leads easily to surface or bulk oxide formation.

This can be appreciated from the CVs in Fig. 38 [138] which show that on Pt in acidic solutions, oxygen reduction takes place at 0.8 V (RHE), regardless of the amount of oxide formed and that, therefore, there is an overpotential of

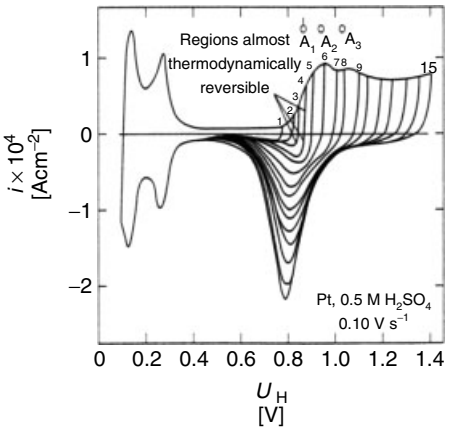


Fig. 38 Platinum electrode voltammograms showing the reversible and irreversible regimes for the formation of platinum oxide species. (0.5 M H_2SO_4, 100 mV s^{-1}, U_H = RHE). (Reprinted with permission from Ref. 138, Copyright 1973 by Elsevier Science).

0.45 V from the thermodynamic potential of 1.23 V for the O_2(g) reduction reaction.

As shown by the volcano plots of Figs 25 and 39 [139], Pt and Pt–Ru alloys are among the most active electrocatalysts for the oxygen evolution reaction. However, RuO_2, which is also the main component of dimensionally stable anode (DSA) electrodes, is significantly more active (Fig. 25).

The nature of O adatoms on noble metals, which frequently form islands on the electrocatalytic surface (Fig. 40), has been studied in detail and imaged by scanning tunneling microscopy (STM), by Ertl and coworkers [140]. Some of the best binary and tertiary alloy electrocatalysts developed so far for the ORR reaction are compared in Fig. 41 [141].

2.7 Oxidation of Small Organic Molecules

The oxidation of small organic molecules, recently reviewed by Iwasita [142], has attracted wide attention during the last 20 years in connection with the development of direct methanol fuel cells (DMFCs)

Table 1 shows the equilibrium oxidation potentials for H_2, CH_3OH, H_2CO, HCOOH, and CO. One observes that the open-circuit potentials of the corresponding fuel cells utilizing these molecules as the fuel at the anode are close to each other and thus show similar promise from a thermodynamic point of view [142]. However, as shown in Fig. 22, the oxidation of CO exhibits an overpotential of 0.7 V on Pt and 0.6 V on Ru anodes in acidic environments. The overpotential for CO oxidation is significantly smaller in

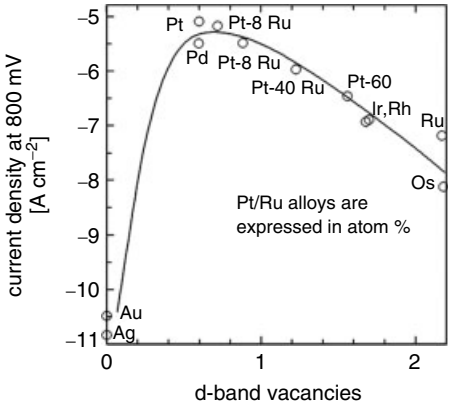


Fig. 39 Volcano plot for oxygen reduction versus d-orbital vacancies. (The d-orbital vacancy for Os is considered to be the same as that for Ru) (85% H_3PO_4 and 25 °C; 900 mV versus RHE corresponds to $\eta = -360$ mV) (Reprinted with permission from Ref. 139, Copyright 1995 by Marcel Dekker, Inc).

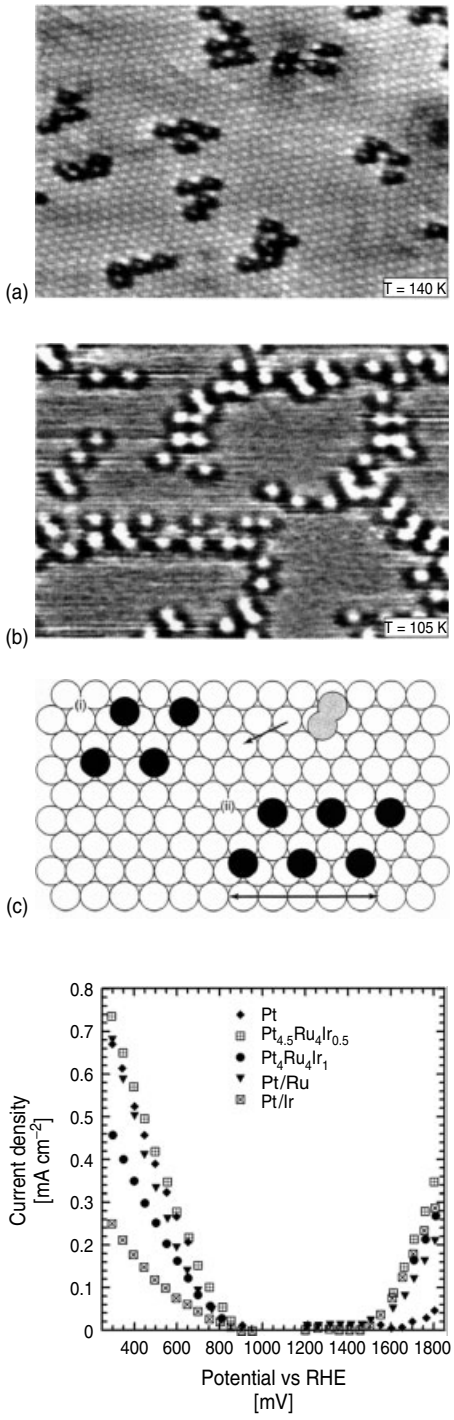


Fig. 40 Atomic-resolution images of oxygen islands on Pt(111). These images illustrate the enhanced precursor dissociation probability at adsorbed oxygen atoms and the resulting quasi-one-dimensional growth for (a) $T = 140$ K (3 L dose, image 100×60 Å^2); (b) $T = 105$ K with chainlike islands corresponding to strings of oxygen pairs (1 L dose, image 70×40 Å^2); (c) model illustrating the directional growth of O_{ad} (black spheres) chains for Pt(111); (i) an oxygen precursor molecule (shaded) approaches an atom island; (ii) on its dissociation a chain segment evolves. (Reprinted with permission from Ref. 140, Copyright 1997 by Nature Publishing Group).

Fig. 41 Combined polarization curves of various catalysts for both oxygen evolution and oxygen reduction. (Reprinted with permission from Ref. 141, Copyright 2001 by Elsevier Science).

alkaline environments [117] but such environments are of no interest for polymer electrolyte membrane fuel cell (PEMFC), applications. The same trend is observed when examining the oxidation of small organic molecules and comparing the performance of Pt electrocatalysts in acidic and alkaline environments, as shown earlier in Fig. 17. The same figure showed the multipeak CVs obtained primarily in acidic environments resulting from the oxidation of passivating oxide and subsequent easier oxidation on the cleaned surface in the reverse scan.

Since such high anodic overpotentials are prohibitive for practical applications, a very intense research effort has been launched during the last 20 years for finding more efficient anodic electrocatalysts.

The source of the problem is that the direct anodic six-electron oxidation

Tab. 1 Thermodynamic data on the oxidation of CO, methanol, and related substances at 25 °C and 1 atm [142]

Theoretical cell reaction	ΔH° [kcal mol^{-1}]	ΔS° [cal mol^{-1}]	ΔG° [kcal mol^{-1}]	n	ΔU°[V]
$H_2 + O_2 \rightarrow H_2O(l)$	−68.14	−39.0	−56.69	2	1.23
$C + 1/2O_2 \rightarrow CO$	−26.4	+21.4	−32.81	2	0.71
$CO + 1/2O_2 \rightarrow CO_2$	−67.62	−20.7	−61.45	2	1.33
$CH_3OH + 3/2O_2 \rightarrow CO_2 + 2H_2O$	−173.8	−23.5	−166.8	6	1.21
$CH_2O + O_2 \rightarrow CO_2 + H_2O$	−134.28	−32.1	−124.7	4	1.35
$HCOOH + 1/2O_2 \rightarrow CO_2 + H_2O$	−64.66	+11.8	−68.2	2	1.48

$$CH_3OH + H_2O \longrightarrow CO_2 + 6H^+ + 6e^- \quad (75)$$

does not take place on any known electrocatalyst, but, as shown by Breiter [143], proceeds via a sequence of steps, which can be generally expressed as

$$CH_3OH \longrightarrow \text{(adsorbed intermediates)} \begin{matrix} \nearrow CO(ad) \\ \searrow H_2CO, HCOOH \rightarrow CO_2 \end{matrix} \quad (76)$$

The formation of adsorbed CO is recognized as a major factor for the observed high anodic overpotentials [142]. In aqueous environments, a good electrocatalyst should be able to

(a) dissociate the C−H bond to form adsorbed methanol;
(b) activate the reaction of H_2O with the resulting carbonaceous intermediate.

On Pt electrodes, which are known to be good catalysts for C−H bond cleavage, CH_3OH cannot displace adsorbed H atoms, so CH_3OH adsorption starts above 0.2 V versus RHE. The second process (b) requires dissociation of H_2O which acts as the oxygen donor of the reaction. On pure Pt, this is possible only at potentials above 0.4–0.45 versus RHE [142]. Thus, on pure Pt, the oxidation of CH_3OH to CO_2 cannot start below 0.45 V and, owing to the inefficient catalysis of the second step, practically starts at 0.7 V, that is, at potentials of no technological interest. The use of on-line mass spectrometry, direct electrochemical mass spectrometry, (DEMS), has revealed several interesting features of the CH_3OH oxidation on Pt

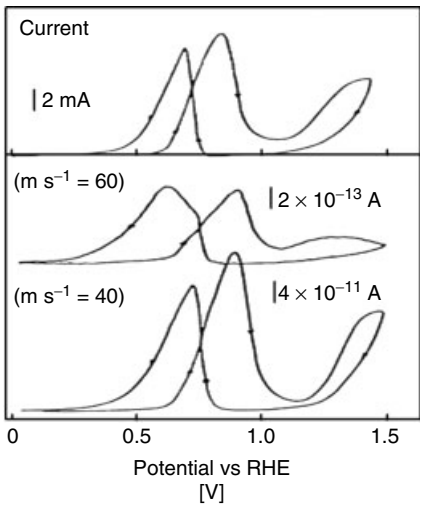


Fig. 42 Direct electrochemical mass spectrometric (DEMS) experiment: current and mass-signal ($m/e = 44$, CO_2 and $m/e = 60$, methyl formate) voltammograms during methanol oxidation on a porous Pt electrode. 0.1 M CH_3OH/1 M $HClO_4$; 20 mV s^{-1}. Surface roughness: ca 50. (Reprinted with permission from Ref. 144, Copyright 1990 by Wiley-VCH).

(Fig. 42) [144]. In addition to CO_2 ($m/e = 44$), significant amounts of methyl formate ($m/e = 60$) are also formed.

Similar direct electrochemical mass spectrometric (DEMS) studies by Vielstich and Anastasijevic [8, 145, 146] of the anodic oxidation of H_2CO in alkaline solutions led to some of the first reports of electrochemical promotion (NEMCA effect) in aqueous solutions, that is, to current efficiencies for CO_2 formation exceeding 100%, owing to the electrochemical activation of a purely catalytic parallel oxidation step on the Pt catalyst electrode [8, 145, 146].

The continuing intense effort for developing efficient anodes for DMFCs usually involves the use of alloys, such as Pt/Ru, where a pronounced synergistic effect is observed (Fig. 43) [11, 147], probably primarily due to bifunctionality. The anode performance is significantly enhanced by the use of various promoting adatoms, such as Bi, which decrease by more than 2 eV the work function of Pt [148] and also Sb and Se where again a pronounced current density enhancement is observed at intermediate adatom coverage values (Fig. 44) [148, 149]. Of particular promise are electrocatalysts involving Pt/Sn where activity is again optimized for θ_{Sn}^{Pt} values close to 0.5 for the electrooxidation of CO, HCOOH, H_2CO, and CH_3OH (Fig. 45) [150]. The use of Pb, Bi, and Sb adatoms for the oxidation of HCOOH on Pt electrodes also leads to very enhanced anodic performance [104]. The investigation of the mechanism of the anodic oxidation of HCOOH is an area of very active research [151–164]. The oxidation of C_2 molecules is also attracting increasing interest [165–169].

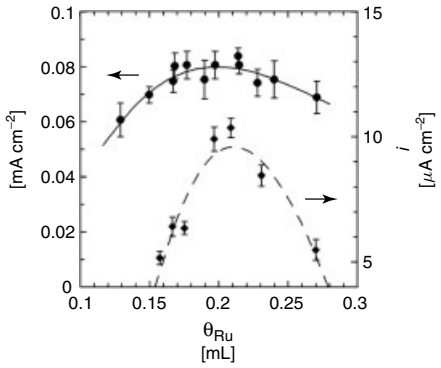


Fig. 43 Methanol electrooxidation current densities as a function of ruthenium coverage at 0.56 V (●) and at 0.38 V (♦) on the Pt(111)/Ru electrode. (Reprinted with permission from Ref. 147, Copyright 1998 by the American Chemical Society).

2.8 Other Important Electrocatalytic Systems

Although the effort for developing efficient PEM, and also solid oxide fuel cell (SOFC), electrocatalysts has absorbed the majority of the research activities in electrocatalysis during the last 15 years, there are also several other theoretically and technologically important electrocatalytic fields that have attracted significant attention and where important advances have been made.

The anodic oxidation of volatile organic components (VOCs) using metal-oxide electrodes has attracted much attention and the group of Comninellis [170–174] has found that often the complete electrochemical oxidation of some organics in aqueous media occurs, without any loss in electrode activity, only at high potentials with concomitant evolution of O_2. Furthermore, it has been found that the nature of electrode material strongly influences both the selectivity and the efficiency of the process.

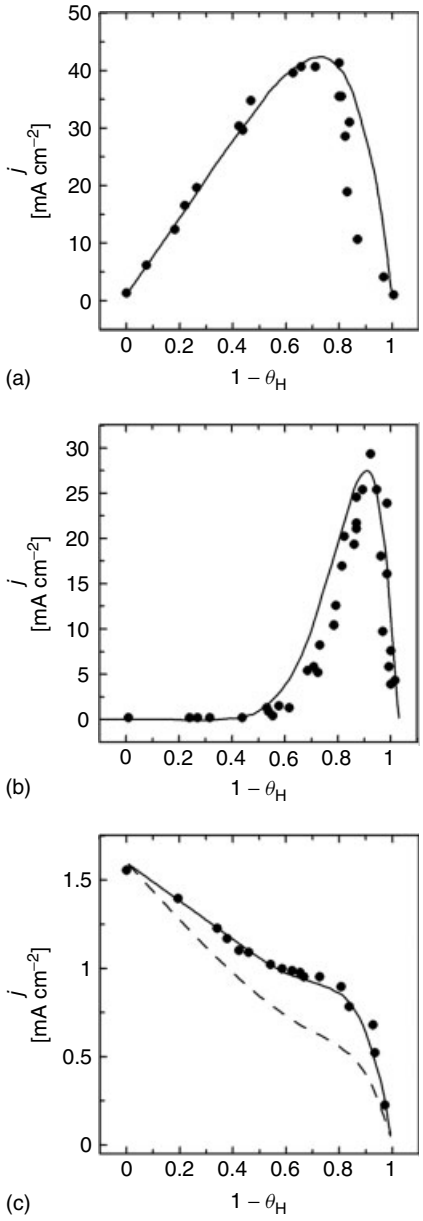


Fig. 44 Comparison between experimental (●) and theoretical currents for formic acid oxidation versus the fraction of platinum surface blocked by the adatom on three different electrodes: (a) Bi/Pt(111), (b) Sb/Pt(111), and (c) Se/Pt(111). (Reprinted with permission from Refs 148, 149, Copyright 1997 by the American Chemical Society).

In order to interpret these observations, a comprehensive model for anodic oxidation of organics in acidic media, including competition with oxygen evolution, has been proposed [170]. This model permits the distinction between the two limiting cases of "active" and "nonactive" anodes. Figure 46 illustrates the reaction scheme in acid medium (e.g., $HClO_4$), where M designates an active site at the anode surface. In all cases, the initial step is the discharge of water molecules to form adsorbed hydroxyl radicals (77):

$$M + H_2O \longrightarrow (M - OH) + H^+ + e^- \quad (77)$$

The electrochemical and chemical reactivities of the adsorbed hydroxyl radicals strongly depend on the nature of electrode material used. Two extreme classes of electrodes can be defined: "active" and "nonactive" electrodes:

At "active" electrodes there is a strong electrode (M)–hydroxyl radical ($^{\bullet}OH$) interaction. In this case, the adsorbed hydroxyl radicals may interact with the anode with possible transition of the oxygen from the hydroxyl radical to the anode surface, forming the so-called higher oxide (Eq. 78). This may be the case when higher oxidation states on the surface electrode are available above the thermodynamic potential for oxygen evolution (1.23 V/RHE).

$$M(^{\bullet}OH) \longrightarrow MO + H^+ + e^- \quad (78)$$

The surface redox couple MO/M can act as a mediator in the oxidation of organics at "active" electrodes (Eq. 79). This reaction competes with the side reaction of oxygen evolution due to the

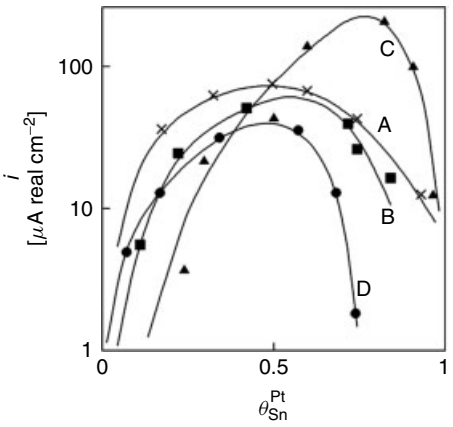


Fig. 45 Polarization curves for the oxidation of CO (A), HCOOH (B), HCHO (C), and CH_3OH (D) at 40 °C on Pt with deposited Sn atoms. (Reprinted with permission from Ref. 150, Copyright 1976 by Elsevier Science).

chemical decomposition of the higher oxide (Eq. 80).

$$MO + R \longrightarrow M + RO \qquad (79)$$

$$MO \longrightarrow M + 1/2O_2 \qquad (80)$$

The oxidation reaction via the surface redox couple MO/M (Eq. 79) may result in the partial (selective) oxidation of organics. At ''nonactive'' electrodes there is a weak electrode (M)–hydroxyl radical ($^{\bullet}OH$) interaction. In this case, the oxidation of organics is mediated by hydroxyl radicals (Eq. 81), which may result in fully oxidized reaction products, that is, CO_2.

$$M(^{\bullet}OH) + R \longrightarrow M + mCO_2 + nH_2O + H^+ + e^- \qquad (81)$$

In this schematic reaction equation, R is the fraction of an organic compound, which contains no heteroatom and needs one atom of oxygen to be transformed to fully oxidized elements. (Actual values of m and n depend on the elemental composition of R to be oxidized.) This reaction competes with the side reaction of hydroxyl radical discharge (direct or indirect through formation of H_2O_2 as intermediate) to O_2 (Eq. 82) without any participation of the anode surface:

$$M(^{\bullet}OH) \longrightarrow M + 1/2O_2 + H^+ + e^- \qquad (82)$$

The distinction between ''active'' and ''nonactive'' behavior and the underlying mechanistic explanation are supported by several experimental observations, including measurement of the concentration of reactive intermediates in the oxygen evolution reaction, such as hydroxyl radicals produced by discharge of water. In fact, hydroxyl radicals produced at ''nonactive'' electrodes can be intercepted by using *p*-nitrosodimethylaniline as a selective scavenger [170].

In practice, however, most anodes will exhibit a mixed behavior since both parallel reaction paths will participate in organics oxidation and oxygen evolution reactions. As a general rule, the closer the reversible potential of the surface redox couple to the potential of oxygen evolution, the higher the active character of the anode. In Ref. 170, dimensionally stable–type anodes (DSA) based on IrO_2 (Ti/IrO_2) and synthetic boron-doped diamond (*p*-Si/BDD) electrodes are compared in the oxidation of organics in acid media (1 M H_2SO_4 or 1 M $HC1O_4$). Ti/IrO_2 may serve as a typical example of ''active'' anode (the redox potential of the IrO_3/IrO_2 couple, 1.35 V/RHE, is close to the standard potential of O_2 evolution, 1.23 V/RHE) and is thus suitable for electrosynthesis, while *p*-Si/BDD, a typical ''nonactive'' BDD, is known to be inert and to have weak adsorption properties.

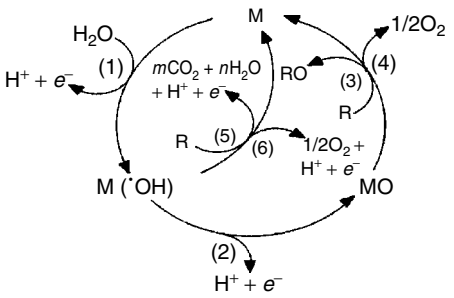


Fig. 46 Scheme of the electrochemical oxidation of organic compounds on ''active'' anodes (Reactions 1–4) and on '''nonactive'' anodes (Reactions 1, 5, 6). (1) Water discharge to hydroxyl radicals, (2) formation of the higher metal oxide, (3) partial (selective) oxidation of the organic compound, R, via the higher metal oxide, (4) oxygen evolution by chemical decomposition of the higher metal oxide, (5) combustion of the organic compound via hydroxyl radicals, and (6) oxygen evolution by electrochemical oxidation of hydroxyl radicals. (Reprinted with permission from Ref. 172, Copyright 1999 by The Electrochemical Society, Inc).

Significant advances have also been made in developing novel materials for oxygen reduction. For example, the use of Chevrel phases and chalcogenides as ORR electrocatalysts in aqueous electrochemistry has been significantly advanced since their discovery in 1986 by Alonso-Vante and Tributsch [175–177].

Recent advances using macrocycles, such as supported metal–phthalocyamine and metal–tetraphenyl-porphyrine, pioneered by Jasinski [178] and pursued by Zagal and others [179–181] show significant promise, although significant optimization is necessary to reach practical current densities.

Similar efforts in solid-state electrochemistry for SOFC development focus on the exploration of new perovskites not only for the ORR but also for the anodic oxidation of hydrocarbons [182]. In this area, the discovery that Cu-based anodes present a viable alternative to the classical Ni-YSZ cermet anodes is particularly noteworthy [166, 183, 184], owing to the significant enhancement of performance by avoiding coke deposition. Similar important advances have occurred in the molten carbonate fuel cell (MCFC) area [9].

The electrocatalysis of hydrogenation reactions has also witnessed significant progress in recent years, particularly the area of the technologically extremely important CO_2 hydrogenation [185–194].

Electrochemical capacitors and supercapacitors have also received increasing attention both for energy storage and also due to the fact that such carbon-supported electrode capacitors play an important role in PEM fuel cell operation [9].

2.9 Electrochemical Promotion of Catalytic Reactions

The phenomenon of EPOC or NEMCA effect was first reported in solid electrolyte systems [23, 195–205], but several NEMCA studies already exist using aqueous electrolyte systems [23, 30, 31, 145] or Nafion membranes [23]. The EPOC phenomenon leads to apparent Faradaic efficiencies, Λ, well in excess of 100% (values up to 10^5 have been measured in solid-state electrochemistry and up to 10^2 in aqueous electrochemistry). This is due to the fact that, as shown by a variety of surface science and electrochemical techniques [23, 40, 195–198, 206–209], the NEMCA effect is due to electrocatalytic (Faradaic) introduction of promoting species onto catalyst-electrode surfaces [23, 196], each of these promoting species being able to catalyze numerous (Λ) catalytic turnovers.

The main experimental features and theory of the EPOC have been reviewed recently [23, 196, 199, 206, 207] and summarized lucidly by Sanchez and Leiva [208, 209].

The basic phenomenology of EPOC when using O^{2-} - or H^{+}-conducting supports is given in Figs 47 and 48 respectively. In the former case (Fig. 47, [23]), the (usually porous) metal (Pt) catalyst

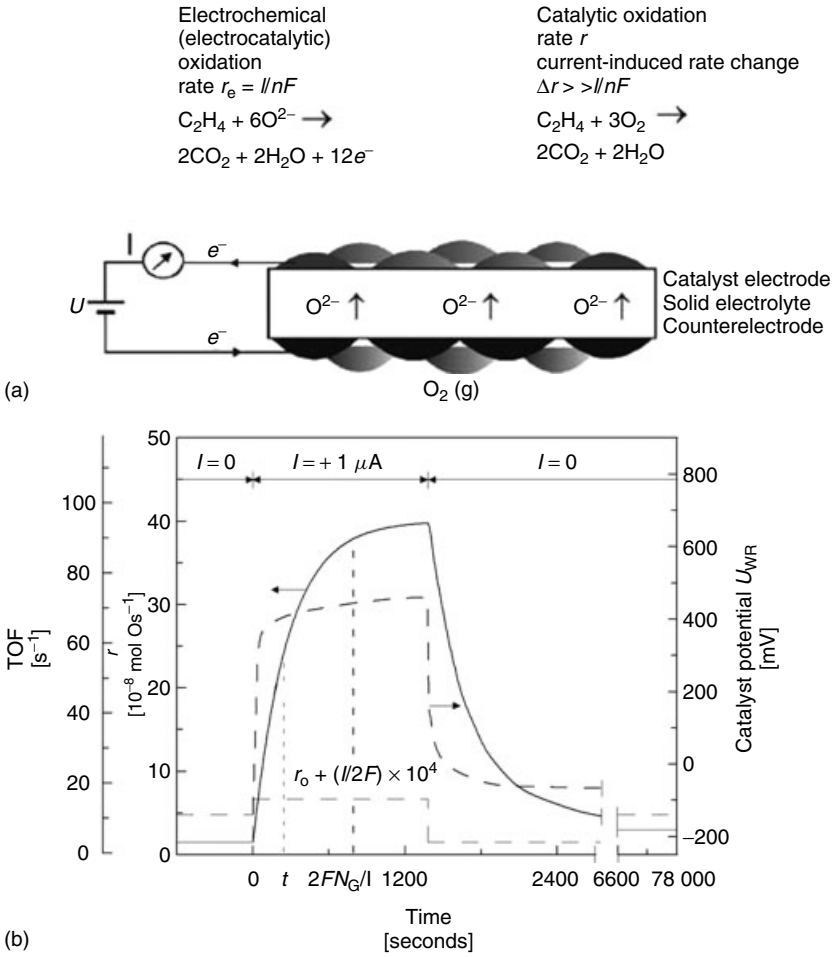


Fig. 47 (a) Basic experimental setup and operating principle of electrochemical promotion with O^{2-} -conducting supports. (b) Catalytic rate, r, and turnover frequency, TOF, response of C_2H_4 oxidation on Pt deposited on YSZ, an O^{2-} conductor, upon step changes in applied current. $T = 370\,^{\circ}C$, $p_{O2} = 4.6$ kPa, $p_{C_2H_4} = 0.36$ kPa. Also shown (dashed line) is the catalyst-electrode potential, U_{WR}, response with respect to the reference, R, electrode. The catalytic rate increase, Δr, is 25 times larger than the rate, r_0, before current application and 74 000 times larger than the rate, $I/2F$, of O^{2-} supply to the catalyst electrode. N_G is the Pt/gas interface surface area, in moles Pt and TOF is the catalytic turnover frequency (moles of O reacting per surface Pt moles per second). (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

electrode, typically 40 nm to 4 µm thick, is deposited on an 8 mol% YSZ solid electrolyte. In the latter case (Fig. 48, Ref. 201) the nanodispersed Pd catalyst is deposited on porous conductive graphitic C which is supported on Nafion, a H^+ conductor.

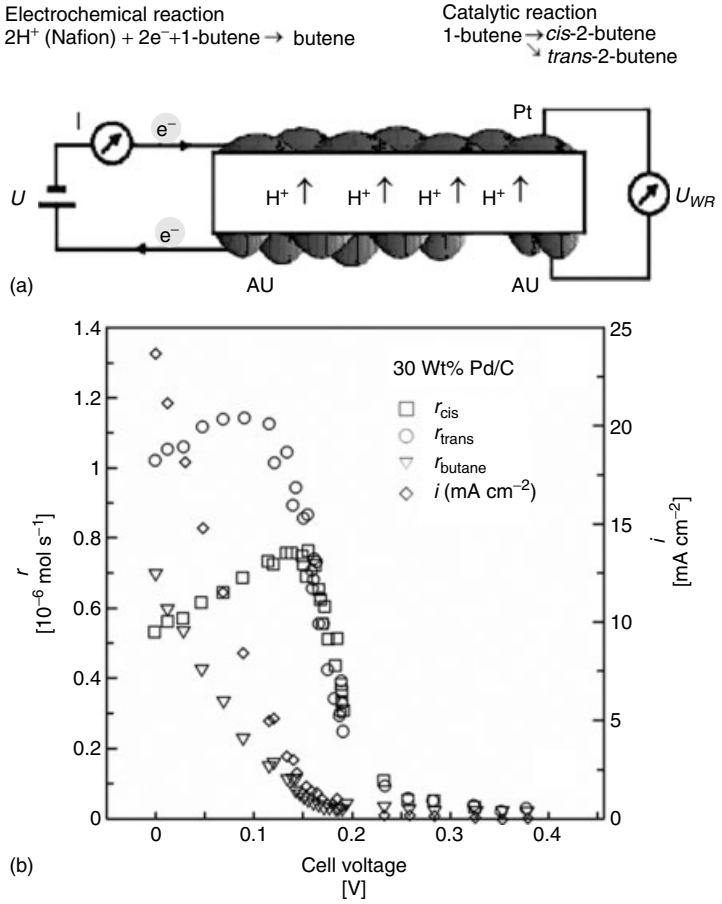


Fig. 48 (a) Basic experimental setup and operating principle of electrochemical promotion using a H^+ conductor during 1-butene isomerization to *cis*-2-butene and *trans*-2-butene on C-supported Pd [23]. (b) Electrochemical promotion of the isomerization reaction [201]. Steady-state effect of cell potential on the current density, and on the rates of formation of *cis*-2-butene, *trans*-2-butene, and butane produced from 1-butene supplied over a dispersed Pd/C catalyst electrode deposited on Nafion, a H^+ conductor [201]. (c) Corresponding effect of cell potential on the selectivities to *cis*-2-butene, *trans*-2-butene, and butane and on the apparent Faradaic efficiency, Λ (filled symbols) defined as $\Delta r_{total}/(I/F)$. Thus, each proton catalyzes the isomerization of up to 50 molecules of 1-butene to *cis*- and *trans*-2-butene. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers (a) and from Ref. 201, Copyright 1997 by the American Chemical Society (b).)

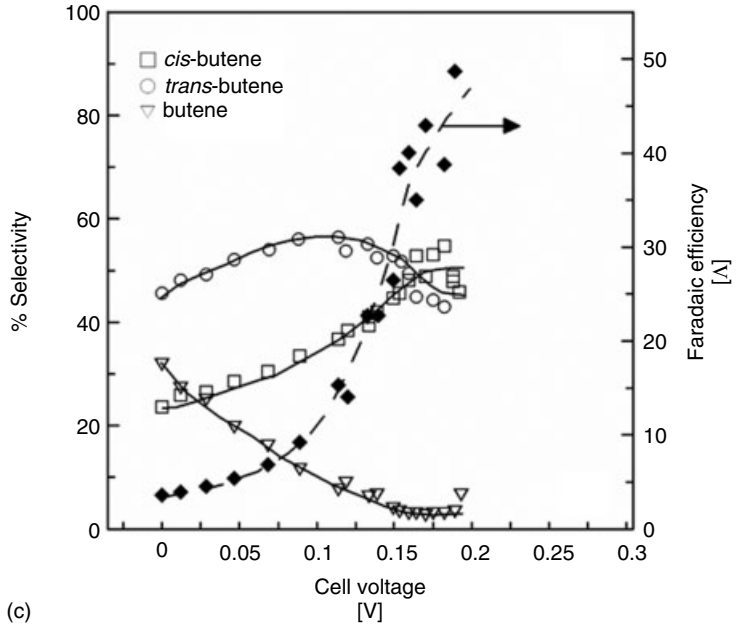


Fig. 48 *Continued.*

In both cases, under open-circuit operation ($I = 0$, no electrochemical rate), there is a catalytic rate, r_0, of ethylene consumption for oxidation to CO_2 (Fig. 47) or of 1-butene consumption owing to reduction to butane and isomerization to *cis*-2-butene and *trans*-2-butene (Fig. 47).

Application of an electrical current, I, or potential (U_{WR}) between the catalyst and a counterelectrode and, thus, changing the catalyst potential, U_{WR}, with respect to a reference electrode, causes very pronounced and strongly non-Faradaic (i.e., $\Delta r \gg I/2F$ in Fig. 47, $\Delta r \gg |-I/F|$ in Fig. 48) alterations to the catalytic rate (Figs 47 and 48) and to the product selectivity (Fig. 48).

Two parameters are commonly used to describe the magnitude of electrochemical promotion: first, the apparent Faradaic efficiency, Λ, defined from

$$\Lambda = \Delta r_{\text{catalytic}}/(I/2F) \qquad (83)$$

where $\Delta r_{\text{catalytic}}$ is the current- or potential-induced change in catalytic rate, I is the applied current, and F is Faraday's constant.

Second, the rate enhancement, ρ, defined from

$$\rho = r/r_o \qquad (84)$$

where r is the electropromoted catalytic rate and r_o is the unpromoted (open circuit) catalytic rate.

A reaction exhibits electrochemical promotion when $|\Lambda| > 1$, whereas electrocatalysis is limited when $|\Lambda| \leq 1$. A reaction is termed *electrophobic* when $\Lambda > 1$ and electrophilic when $\Lambda < -1$. In the former case, the rate increases with catalyst potential, U, whereas in the latter case the rate decreases with catalyst potential. Λ values up to 3×10^5 [23, 210] and ρ values up to 150 [23] have been found for several systems. More recently, ρ values between 300 and 1200 [211, 212] have been measured for C_2H_4 oxidation on Pt.

In the experiment shown in Fig. 47, it is $\Lambda = 74\,000$ and $\rho = 26$, that is, the rate of C_2H_4 oxidation increases by a factor of 25 and the increase in the rate of O consumption is 74 000 times larger than the rate, $I/2F$, of O^{2-} supply to the catalyst. In the experiment shown in Fig. 48, the maximum ρ values for the production of *cis*-2-butene, *trans*-2-butene, and butene are of the order of 50 and the corresponding maximum Λ values are of the order of 40 for *cis*-2-butene formation, 10 for *trans*-2-butene formation, and less than 1 for butene formation. Thus, each proton supplied to the Pd catalyst can cause the isomerization of up to 40 1-butene molecules to *cis*-2-butene and up to 10 1-butene molecules to *trans*-2-butene, whereas the hydrogenation of 1-butene to butane is electrocatalytic, that is, Faradaic.

So far, more than 70 different catalytic reactions (oxidations, hydrogenations, dehydrogenations, isomerizations, decompositions) have been electrochemically promoted on Pt, Pd, Rh, Ag, Au, Ni, IrO_2, RuO_2 catalysts deposited on O^{2-} (YSZ), Na^+ (β''-Al_2O_3), H^+ ($CaZr_{0.9}In_{0.1}O_{3-\alpha}$, Nafion), F^- (CaF_2), aqueous, molten salt, and mixed ionic-electronic (TiO_2, CeO_2) conductors [23].

These studies have shown that, quite often, over relatively wide (e.g., 0.3–0.5 V) ranges of potential, the catalytic rates depend on catalyst-electrode potential in an exponential manner, similar to the high field approximation of the Butler–Volmer equation, that is,

$$r/r_o = \exp\left(\frac{\alpha e \Delta U}{k_b T}\right) = \exp\left(\frac{\alpha \Delta \Phi}{k_b T}\right) \tag{85}$$

where r_o is the unpromoted (i.e., open-circuit) catalytic rate, ΔU is the applied overpotential, $\Delta\Phi$ is the overpotential-induced change in the catalyst-electrode work function and α (typically $|\alpha| \approx 0.2-1$) is a parameter that is positive for electrophobic reactions ($\partial r/\partial\Phi > 0$, $\Lambda > 1$) and negative for electrophilic reactions ($\partial r/\partial\Phi < 0$, $\Lambda < -1$) (Fig. 53).

The second equality (85) holds because, as shown by Kelvin probe [23, 195] and ultra-violet electron spectroscopic (UPS), [23] work function measurements, the equality

$$e\Delta U_{WR} = \Delta\Phi \tag{86}$$

is valid over wide (e.g., 0.5–1 V) catalyst-electrode potential ranges in solid-state electrochemistry. This equation is identical to that reported in aqueous electrochemistry for emersed electrodes [213, 214].

Equation (86) is a limiting case of the general equation

$$e\Delta U_{WR} = -\Delta\overline{\mu}(= -\Delta E_F) = \Delta\Phi + e\Delta\Psi \tag{87}$$

valid for any electrochemical cell [23], where $\overline{\mu}$ is the electrochemical potential of electrons in the catalyst electrode, $E_F(=\overline{\mu})$ is the Fermi level of the catalyst-electrode, and Ψ is the outer (Volta) potential of the metal catalyst electrode in the gas outside the metal/gas interface. The latter vanishes ($\Psi = 0$, $\Delta\Phi = 0$) when no net charge resides at the metal/gas interface [23, 215]. Thus, the experimental Eq. (86) manifests the formation of a neutral double layer, termed *effective* double layer, at the metal/gas interface (Fig. 49). At the molecular level, the stability of the EDL and thus the validity of Eq. (86) requires that the migration (backspillover) of the promoting ion ($O^{\delta-}$, $Na^{\delta+}$) is fast relative to its desorption or catalytic consumption. When this condition is not met (e.g., high T or nonporous electrodes) or also when the limits of zero or saturation coverage of the promoting ion

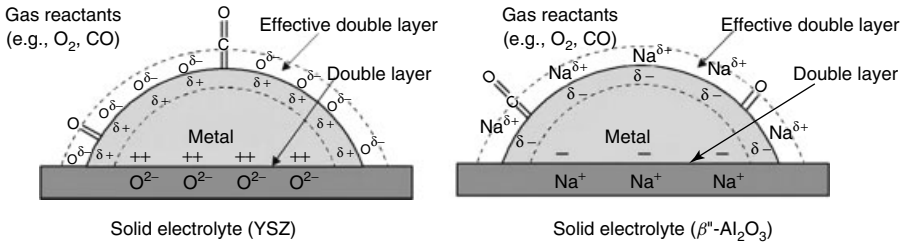


Fig. 49 Schematic representation of a metal catalyst electrode deposited on a O^{2-} - and a Na^+-conducting solid electrolyte, showing the location of the metal–solid electrolyte double layer and that of the effective double layer created at the metal/gas interface due to potential-controlled ion migration (backspillover). (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

are reached (at very positive or negative ΔU_{WR}), then deviations from Eq. (1) are observed [23, 215].

From Eq. (86) it follows that in electrochemical promotion the work function of the catalyst surface can be in situ controlled by the applied potential U_{WR} and thus electrochemical promotion is, simply, catalysis in presence of an electrochemically controllable double layer at the catalyst/gas interface [23]. The EDL affects the binding strength of chemisorbed reactants and reaction intermediates and thus affects the catalytic rate in a very pronounced and reversible manner.

The molecular origin of electrochemical promotion is currently understood on the basis of the sacrificial promoter mechanism [23]. NEMCA results from the Faradaic (i.e., at a rate I/nF) introduction of promoting species ($O^{\delta-}$ in the case of O^{2-} conductors, H^+ in the case of H^+ conductors) on the catalyst surface. This electrochemically introduced O^{2-} species acts as a promoter for the catalytic reaction (by changing the catalyst work function and affecting the chemisorptive bond strengths of coadsorbed reactants and intermediates) and is eventually consumed at a rate equal, at steady state, to its rate of supply ($I/2F$) which is Λ times smaller than the rate of consumption of the catalytic reactant, for example, atomic O originating from the gas phase [23].

Figure 50 shows the validity of the sacrificial promoter concept for the galvanostatic transient in Fig. 47, by presenting O_2 TPD (Fig. 50b) and cyclic voltammetric (Fig. 50c) spectra obtained at times corresponding to those of the NEMCA galvanostatic transient of C_2H_4 oxidation (Fig. 47), under high vacuum conditions. One clearly observes, both with TPD and with cyclic voltammetry, the Faradaic introduction, over a time period $2FN_G/I$ (where N_G is the catalyst surface area expressed in moles and thus $2FN_G/I$ is the time required to form a monolayer of O^{2-} on the catalyst surface), of a second (backspillover) strongly bonded oxygen species on the Pt catalyst surface which displaces the normally chemisorbed oxygen state to lower desorption temperatures. This displacement, which results from strong repulsive lateral interactions, between O^{2-} and more covalently bonded atomic oxygen [23] causes the observed pronounced enhancement in the catalytic rate. The backspillover O^{2-} state acts as a sacrificial promoter. This molecular picture has been recently confirmed by the use of $^{18}O_2$ TPD [41].

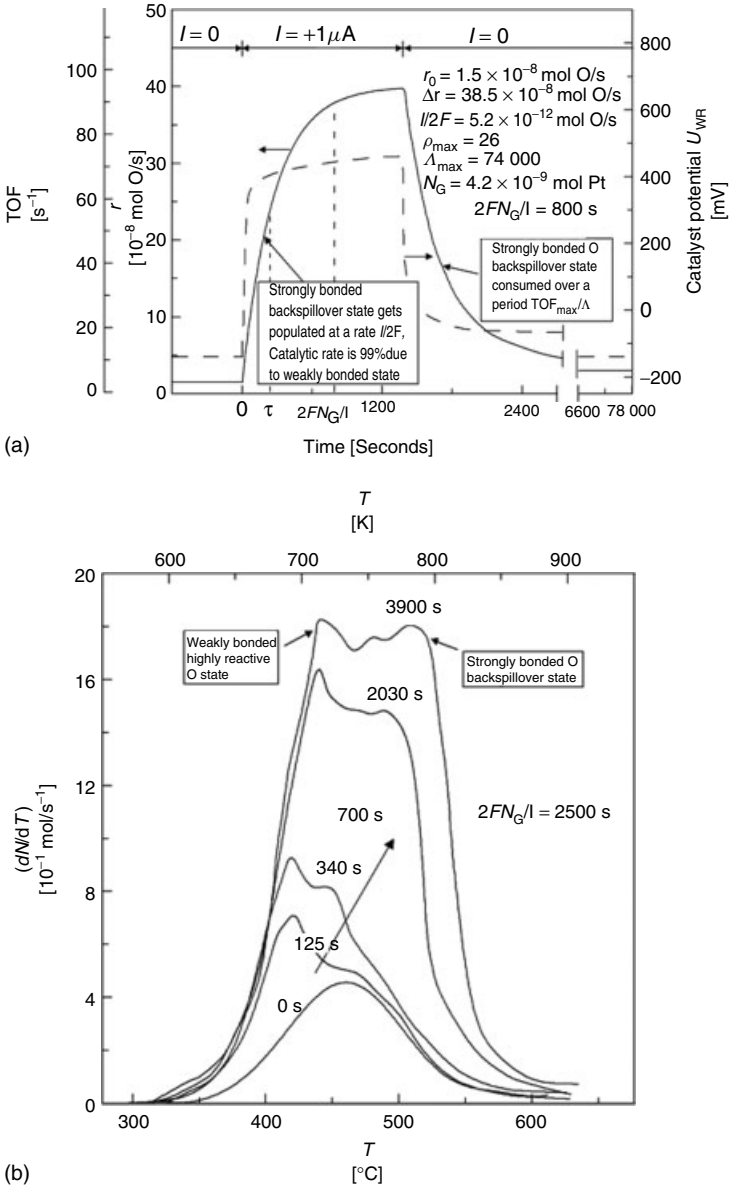


Fig. 50 NEMCA and its origin on Pt/YSZ catalyst electrodes [23]. Transient effect of the application of a constant current (a,b) or constant potential U_{WR} (c) on (a) the rate, r, of C_2H_4 oxidation on Pt/YSZ (also showing the corresponding U_{WR} transient), (b) the O_2 TPD spectrum on Pt/YSZ after current ($I = 15$ μA) application for various times t, and (c) the cyclic voltammogram of Pt/YSZ after holding the potential at $U_{WR} = 0.8$ V for various times t. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

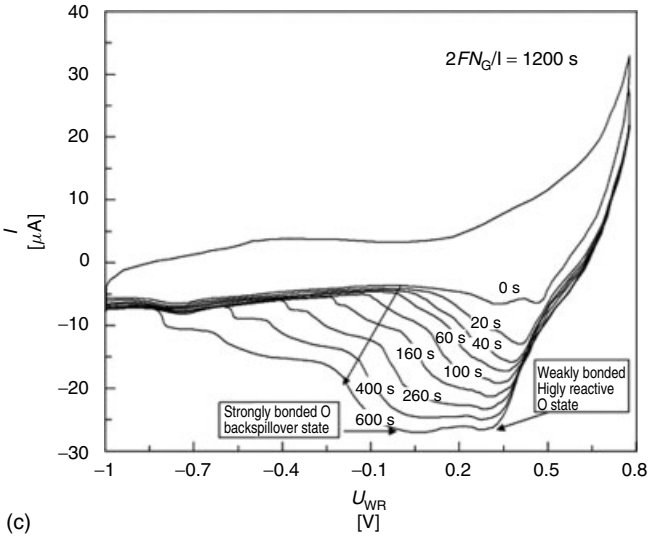


Fig. 50 *Continued.*

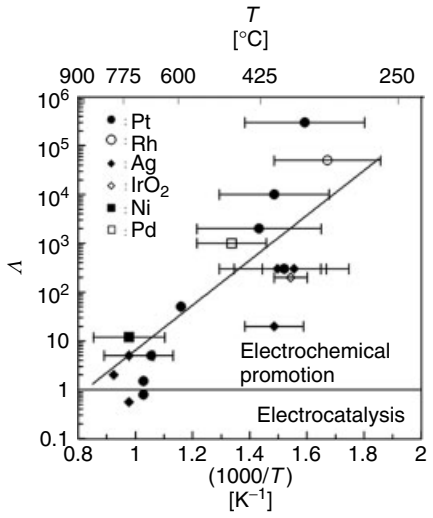


Fig. 51 Effect of temperature on the Faradaic efficiency, Λ, values measured in electrochemical promotion (NEMCA) studies of C_2H_4 oxidation on various metals [23]. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

The limits of electrocatalysis ($|\Lambda| \leq 1$) and electrochemical promotion ($|\Lambda| > 1$) are defined by the parameter $2Fr_o/I_0$ (Eq. 31, Refs 23, 208) which provides an estimate of $|\Lambda|$. In general, $|\Lambda|$ decreases with temperature until the limit of electrocatalysis and electrochemical promotion are reached as shown in Fig. 51.

2.10 Promotional Rules in Catalytic and Electrocatalytic Systems

In recent years, and starting from electrochemical promotion studies [23], certain rules have been extracted for the selection of promoters for catalytic and electrocatalytic systems.

From Eq. (86), one can study the dependence of catalytic reactions on catalyst work function. Four types of rate-work function dependence have been identified experimentally, that is, electrophobic ($\partial r/\partial\Phi > 0$), electrophilic ($\partial r/\partial\Phi < 0$), volcano (where the rate exhibits a maximum with varying Φ) and inverted volcano type, where the rate exhibits a minimum with varying Φ (Fig. 52).

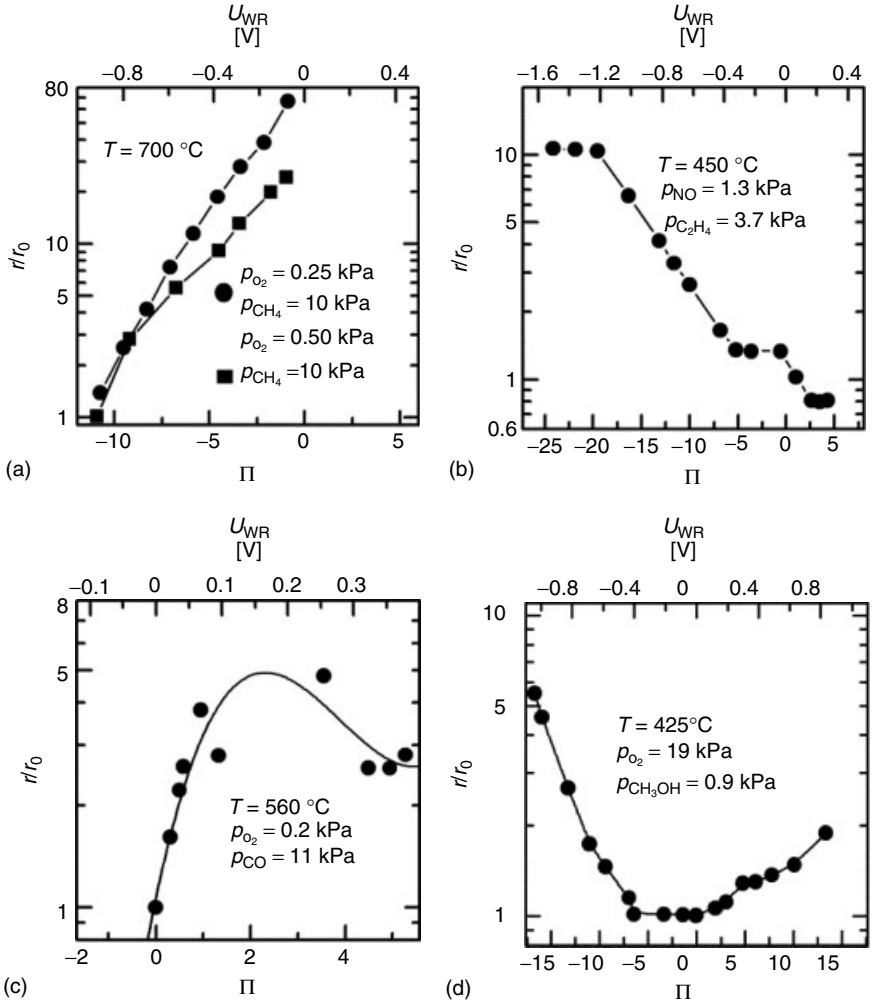


Fig. 52 Examples for the four types of electrochemical promotion behavior: (a) electrophobic, (b) electrophilic, (c) volcano type, and (d) inverted volcano type. (a) Effect of catalyst potential and work function change (versus $I = 0$) on the rate of CH_4 oxidation for high (20:1) and (40:1) CH_4 to O_2 feed ratios, Pt/YSZ. (b) Effect of catalyst potential on the rate enhancement ratio for the rate of NO reduction by C_2H_4 consumption on Pt/YSZ. (c) NEMCA-generated volcano plots during CO oxidation on Pt/YSZ. (d) Effect of dimensionless catalyst potential on the rate constant of H_2CO formation, Pt/YSZ ([23], original references therein). $\Pi = FU_{WR}/RT$ $(= \Delta\Phi/k_bT)$. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

There exist rules which allow one to predict the type to which a given catalytic reaction belongs on the basis of the reaction kinetics with respect to the electron acceptor (A) or electron donor (D) reactants [23, 27]. The latter are defined operationally by the sign of the change induced in catalyst-electrode work

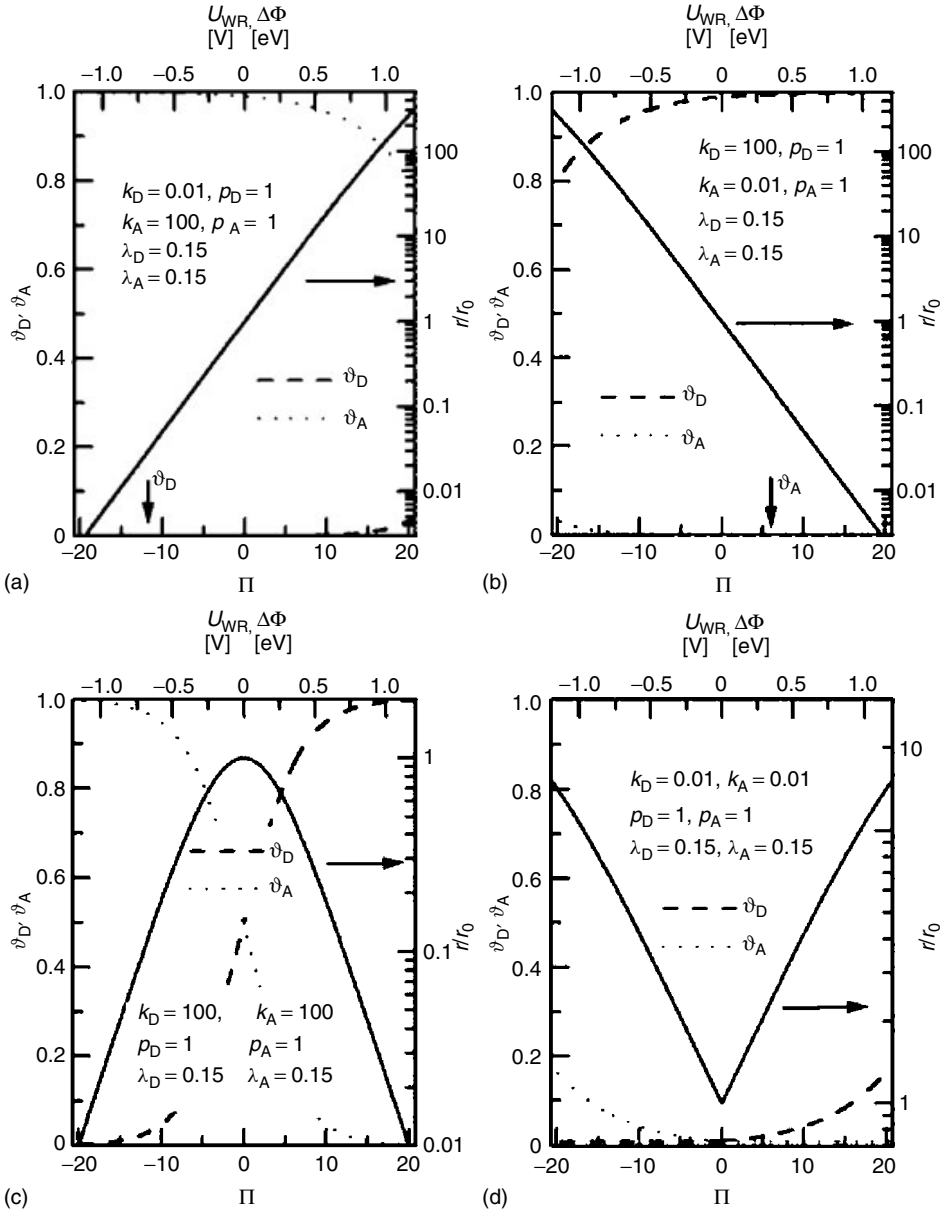


Fig. 53 Effective double-layer model–predicted electrochemical promotion behavior: (a) electrophobic, (b) electrophilic, (c) volcano type, and (d) inverted volcano type [23], showing the dependence of catalytic rate and corresponding coverages of the electron donor (D) and electron acceptor (A) reactants on catalyst-electrode potential, work function, and dimensionless potential $\Pi = eU_{WR}/k_bT$. One observes the transition between the four promotional behaviors upon varying the equilibrium adsorption constants k_A and k_D. (Reprinted with permission from Ref. 23, Copyright 2001 by Kluwer/Plenum Publishers).

function, Φ, by the adsorption of a reactant. In the case of an electron acceptor (A, e.g., O_2), it is $\partial\Phi/\partial\theta_A > 0$, whereas in the case of an electron donor (D, e.g., C_2H_4), it is $\partial\Phi/\partial\theta_D < 0$ [23, 27].

These rules have recently been shown to apply to supported heterogeneous catalysts as well, owing to the thermally induced spillover of ions from the support to the metal/gas interface [23] and concomitant establishment of an EDL at the metal–gas interface [23, 27]. According to these rules, a reaction is electrophobic ($\partial r/\partial\Phi > 0$) when $\partial r/\partial p_D > 0$ and $\partial r/\partial p_A \leq 0$ and electrophilic ($\partial r/\partial\Phi < 0$) when $\partial r/\partial p_D \leq 0$ and $\partial r/\partial p_A > 0$. In the former case the rate, is enhanced with electronegative promoters, whereas in the latter case it is enhanced with electropositive promoters.

Electrophobic and electrophilic catalytic reactions are similar to electrocatalytic oxidations and reduction reactions, respectively, since in the cases of electrophobic catalytic reaction and electrocatalytic oxidations the rate increases with increasing potential, whereas in the case of electrophilic catalytic reactions or electrocatalytic reductions the rate increases with decreasing potential.

One can use the EDL isotherm [23, 27]:

$$k_j p_j = (\theta_j/(1-\theta_j)) \exp(-\lambda_j \Pi) \quad (88)$$

which is mathematically equivalent to the Frumkin isotherm, Eq. (7), where λ_j (± 0.1–0.5) are the partial charge-transfer parameters of the reactants in the adsorbed state [23] and $\Pi = \Delta\Phi/k_bT = \frac{e\Delta U}{k_bT}$, to derive analytical mathematical expressions for catalytic promotional kinetics [23]. Thus, for the case of surface reaction rate control the corresponding expression is:

$$r = \frac{k_R k_A k_D p_A p_D \exp[(\lambda_D + \lambda_A)\Pi]}{[1 + k_D p_D \exp(\lambda_D \Pi) + k_A p_A \exp(\lambda_A \Pi)]^2} \quad (89)$$

where $k_R = k_R^o \exp(\lambda_R \Pi)$ and λ_R is the partial charge-transfer parameter of the transition state. Figure 52 shows plots of Eq. (89) for different values of k_D and k_A and shows the transition from electrophobic to electrophilic to volcano type and to inverted volcano type behavior by simply varying the values of the adsorption equilibrium constants k_D and k_A.

The above rules can also help rationalize some of the key observations in electrocatalytic systems of great theoretical and practical importance. The example of CO oxidation on Pt in aqueous solutions is quite illustrative: It is well established [117] that the activity of Pt(111) increases dramatically in the sequence $Br^- \ll HClO_4 \ll$ NaOH or KOH and the oxidation ignition starts at 1.1, 0.92, and 0.65 V (versus RHE) [117]. In the latter case (0.1 M KOH), the onset of the preignition starts at 0.25 V (versus RHE), that is, in the H_{upd} potential region. The CO oxidation proceeds via reaction between adsorbed CO and OH^- in a Langmuir–Hinshelwood (L–H)-type mechanism. This implies that in alkaline solutions adsorbed OH can exist even at potentials below 0.25 V (versus RHE).

In view of the fact that in this low-potential region the surface is predominantly covered by CO with very small coverage of the electron acceptor surface oxygen species, this L–H reaction is, by definition, electrophilic and is thus expected to be promoted by electron donor species, such as Na^+ present at the OHP, as experimentally observed [117].

Conversely, the presence of adsorbed U^{o}_{WR} of $HClO_4^-$ ions is expected to further hinder the adsorption of the rate-limiting oxygen-containing reactant, further poisoning the reaction, as also experimentally observed [117].

Thus, in view of the fact that in the CO oxidation system in aqueous solutions CO has significantly lower electron acceptor character than coadsorbed oxygen, that is, it behaves operationally as the electron donor reactant which is more strongly bound to the surface, it follows that the addition of electropositive promoters (e.g., NaOH) or the alloying of Pt with metals which lowers its work function, will enhance the rate, in agreement with experiment. These observations strongly underline the close similarity of electrocatalytic and catalytic reactions. Similar to heterogeneous catalysis, metal alloying, such as Pt_3Sn, leads to both ligand (electronic) and bifunctional effects [9, 117] which can act cooperatively to enhance the electrocatalyst performance.

2.11 Concluding Remarks

The urgent need to develop more efficient fuel cell anodes and cathodes has brought the electrochemical, catalytic, and surface science communities closer than ever before and has made electrocatalysis a rapidly growing field both in experimental new findings and in theoretical understanding. It is very likely that the rapid advances in catalyst and electrocatalyst nanoparticle preparation and characterization [10], together with the utilization of new powerful in situ techniques, such as electrochemical nuclear magnetic resonance (NMR) [216] in conjunction with the rapidly developing theoretical and computational methods [89] will soon become valuable tools not only for a better fundamental understanding of electrocatalysis but also for the design of more efficient practical electrocatalysts.

References

1. W. T. Grubb, *17th Annual Power Sources Conference*, Atlantic City, 1963.
2. F. B. Bowden, E. Rideal, *Proc. R. Soc.* **1928**, *A120*, 59–79.
3. (a) I. Horiuti, M. Polanyi, *Acta Physicohimica U.R.S.S.* **1935**, *2*, 505–532; (b) *Trans. Faraday Soc.* **1934**, *30*, 1164–1172.
4. J. O. M. Bockris, Z. S. Minevski, *Electrochim. Acta* **1994**, *39*, 1471–1479.
5. J. O. M. Bockris, A. K. N. Reddy, M. Gamboa-Aldeco, *Modern Electrochemistry*, 2nd ed., Kluwer Academic/Plenum Publishers, New York, 2000
6. J. O. M. Bockris, S. U. M. Khan, *Surface Electrochemistry, A Molecular Level Approach*, Plenum Press, New York, 1993, pp. 319–349, 394–405.
7. A. J. Bard, L. R. Faulkner, *Electrochemical Methods. Fundamentals and Applications*, John Wiley and Sons, New York, 2001.
8. C. H. Hamann, A. Hamnett, W. Vielstich, *Electrochemistry*, 2nd ed., Wiley-VCH, Weinheim, 1998.
9. W. Vielstich, A. Lamm, H. A. Gasteiger, (Eds.), *Handbook of Fuel Cells: Fundamentals Technology and Applications, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
10. A. Wieckowski, E. Savinova, C. G. Vayenas, (Eds.), *Catalysis and Electrocatalysis at Nanoparticles*, Marcel Dekker, New York, 2003.
11. P. N. Ross, in *Electrocatalysis* (Eds.: J. Lipkowski, P. N. Ross), John Wiley and Sons-VCH, New York, 1998, pp. 43–74.
12. R. Parsons, J. O. M. Bockris, *Trans. Faraday Soc.* **1951**, *47*, 914.
13. B. E. Conway, J. O. M. Bockris, *Can. J. Chem.* **1957**, *35*, 1124.
14. P. George, J. Griffith, in *Enzymes* (Eds.: P. D. Boyer, H. Lardy, K. Myrback), Academic Press, New York, 1959, p. 347, Vol. 1.

15. V. Levich, R. R. Dogonadze, *Dokl. Akad. Nauk SSSR* **1959**, *124*, 123.
16. R. A. Marcus, *J. Chem. Phys.* **1965**, *93*, 679.
17. J. O. M. Bockris, S. U. M. Khan, *Quantum Electrochemistry, A Monograph*, Plenum Press, New York, 1979, p. 538.
18. W. Schmickler, *J. Electroanal. Theory* **1979**, *100*, 533.
19. W. Schmickler, *J. Electroanal. Chem.* **1990**, *284*, 269.
20. M. J. Weaver, *Chem. Rev.* **1992**, *92*, 463.
21. W. Schmickler, *Chem. Phys. Lett.* **1995**, *237*, 152.
22. W. Schmickler, *Interfacial Electrochemistry*, Oxford University Press, Oxford, 1996.
23. C. G. Vayenas, S. Bebelis, C. Pliangos et al., *Electrochemical Activation of Catalysis: Promotion, Electrochemical Promotion and Metal-Support Interactions*, Kluwer Academic/Plenum Publishers, New York, 2001.
24. D. Tsiplakides, S. Neophytides, C. G. Vayenas, *Ionics* **1997**, *3*, 201–208.
25. G. Rangelov, L. Surnev, *Surf. Sci.* **1987**, *185*, 457–468.
26. G. Pacchioni, F. Illas, S. Neophytides et al., *J. Phys. Chem.* **1996**, *100*, 16653–16661.
27. S. Brosda, C. G. Vayenas, *J. Catal.* **2002**, *208*, 38–53.
28. A. J. McEvoy, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 394–407, Vol. 2.
29. E. Gileadi, *Electrode Kinetics for Chemists, Chemical Engineers and Material Scientists*, Wiley-VCH, New York, 1993.
30. S. Neophytides, D. Tsiplakides, P. Stonehart et al., *Nature* **1994**, *370*, 45–47.
31. L. Ploense, M. Salazar, B. Gurau et al., *J. Am. Ceram. Soc.* **1997**, *119*, 11550.
32. F. Sapountzi, M. N. Tsampas, C. G. Vayenas, *Top. Catal.* **2006**, in press.
33. P. A. Thiel, T. E. Madey, *Surf. Sci. Rep.* **1987**, *7*, 211.
34. C. Sánchez, E. Leiva, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 4–13, Vol. 2.
35. J. G. Gordon, O. R. Melroy, M. F. Toney, *Electrochim. Acta* **1995**, *40*, 3–8.
36. T. Iwasita, X. H. Xia, H. D. Liess et al., *J. Phys. Chem. B* **1997**, *101*, 7542.
37. K. Ataka, T. Yotsyanagi, M. Osawa, *J. Phys. Chem.* **1996**, *10*, 10664.
38. W. Vielstich, Cyclic voltammetry, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 153–162, Vol. 2.
39. T. Chao, K. J. Walsh, P. S. Fedkiw, *Solid State Ionics* **1991**, *47*, 277–285.
40. S. G. Neophytides, C. G. Vayenas, *J. Phys. Chem.* **1995**, *99*, 17063–17067.
41. A. Katsaounis, Z. Nikopoulou, X. E. Verykios et al., *J. Catal.* **2004**, *222*, 192–206.
42. A. D. Frantzis, S. Bebelis, C. G. Vayenas, *Solid State Ionics* **2000**, *136-137*, 863–872.
43. D. Kek, M. Mogensen, S. Pejovnik, *J. Electrochem. Soc.* **2001**, *148*(8), A878–A886.
44. G. Ertl, H. Knötzinger, J. Weitcamp, *Handbook of Catalysis*, VCH Publishers, Weinheim, 1997.
45. T. J. Schmidt, H. A. Gasteiger, G. D. Staeb et al., *J. Electrochem. Soc.* **1998**, *145*, 2355.
46. H. A. Gasteiger, N. Markovic, P. N. Ross Jr et al., *J. Phys. Chem.* **1994**, *98*, 617.
47. S. Trasatti, Reaction mechanism and rate determining steps, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 79–87, Vol. 2.
48. R. Parsons, The structure of the electrical double layer and its influence on the rates of electrode reactions, in *Advances in Electrochemistry and Electrochemical Engineering* (Eds.: P. Delahay), Wiley-Interscience, New York, 1961, pp. 1–64, Vol. 1.
49. A. N. Frumkin, *Z. Phys. Chem.* **1933**, *164*, 121.
50. B. E. Conway, M. Salomon, *Electrochim. Acta* **1964**, *9*, 1599.
51. S. Trasatti, Electrode kinetics and electrocatalysis of hydrogen and oxygen electrode reactions. 4. The oxygen evolution reaction, in *Electrochemical Hydrogen Technologies* (Ed.: H. Wendt), Elsevier, Amsterdam, 1990, pp. 104–135.
52. (a) B. E. Conway, D. P. Wilkinson, *J. Electroanal. Chem.* **1986**, *210*, 167; (b) **1986**, *214*, 633.
53. M. T. M. Koper, W. Schmickler, A unified model for electron and ion transfer reactions on metal electrodes, in *Electrocatalysis*

(Eds.: J. Lipkowski, P. N. Ross), Wiley-VCH, New York, 1998, pp. 291–322.
54. S. Trasatti, *Electrochim. Acta* **1984**, *29*(11), 1503–1512.
55. B. E. Conway, G. Jerkiewicz, *Electrochim. Acta* **2000**, *45*, 4075–4083.
56. D. Y. Wang, A. S. Nowick, *J. Electrochem. Soc.* **1981**, *128*(1), 55.
57. S. Trasatti, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 88–92, Vol. 2.
58. M. T. M. Koper, G. A. Voth, *J. Chem. Phys.* **1998**, *109*, 1991.
59. A. Wieckowski, (Ed.), *Interfacial Electrochemistry*, Marcel Dekker, New York, 1999, pp. 769–954.
60. R. Parsons, *Trans. Faraday Soc.* **1958**, *54*, 1053.
61. (a) H. Gerischer, *Z. Phys. Chem. NF* **1956**, *8*, 137; (b) *Bull. Soc. Chem. Belg.* **1958**, *67*, 506.
62. (a) L. I. Krishtalik, *Zh. Fiz. Khim.* **1957**, *31*, 2403. (b) **1959**, *33*, 1715.
63. O. A. Petrii, G. A. Tsirlina, *Electrochim. Acta* **1994**, *39*, 1739.
64. S. Trasatti, E. Lust, The potential of zero charge, in *Modern Aspects of Electrochemistry* (Eds.: R. E. White, J. O. M. Bockris, B. E. Conway), Plenum, New York, 1999, pp. 1–215, Vol. 33.
65. S. Trasatti, Interfacial electrochemistry of conductive metal oxides for electrocatalysis, in *Interfacial Electrochemistry: Theory, Experiment, and Applications* (Ed.: A. Wieckowski), Marcel Dekker, New York, 1999, pp. 769–792.
66. C. Sánchez, E. Leiva, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 14–35, Vol. 2.
67. L. I. Schiff, Approximation methods for bound states, in *Quantum Mechanics*, McGraw-Hill, New York, 1968, pp. 279–295.
68. R. A. Marcus, *J. Chem. Phys.* **1965**, *43*, 679.
69. W. Schmickler, in *Interfacial Electrochemistry*, Oxford University Press, New York, 1996, Chapters 3 and 5.
70. (a) H. Gerischer, *Z. Phys. Chem.* **1960**, *F 26*, 223. (b) **1961**, *27*, 40.
71. W. Schmickler, *Ber. Bunsen-Ges. Phys. Chem.* **1978**, *82*, 477.
72. J. Ulstrup, *Surf. Sci.* **1986**, *101*, 564.
73. W. Schmickler, *J. Electroanal. Chem.* **1986**, *204*, 31.
74. W. Schmickler, M. T. M. Koper, *Electrochem. Commun.* **1999**, *1*, 402.
75. M. Koper, W. Schmickler, A unified model for electron and ion transfer reactions on metal electrodes, in *Electrocatalysis* (Eds.: J. Lipkowski, P. N. Ross), Wiley VCH, New York, 1998.
76. M. T. M. Koper, A. P. J. Jansen, R. A. van Santen et al., *J. Chem. Phys.* **1998**, *109*, 6051.
77. H. A. Gasteiger, N. Markovic, P. N. Ross Jr et al., *J. Phys. Chem.* **1994**, *98*, 617.
78. P. N. Ross, in *Electrocatalysis* (Eds.: J. Lipkowski, P. N. Ross), John Wiley and Sons, New York, 1998.
79. E. Shustorovich, H. Sellers, *Surf. Sci. Rep.* **1998**, *31*, 1.
80. H. A. Gasteiger, N. Markovic, P. Ross et al., *J. Phys. Chem.* **1993**, *97*, 12020.
81. D. S. Corrigan, M. J. Weaver, *J. Electroanal. Chem.* **1988**, *241*, 143.
82. F. Jensen, *Introduction to Computational Chemistry*, John Wiley and Sons, Chichester, 1999.
83. P. Hohenberg, W. Kohn, *Phys. Rev., B* **1964**, *136*, 864.
84. M. T. M. Koper, Numerical simulation of electrocatalytic processes, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
85. B. Hammer, J. K. Nørskov, *Adv. Catal.* **2001**, *45*, 71.
86. R. A. van Santen, M. Neurock, *Catal. Rev.–Sci. Eng.* **1995**, *37*, 357.
87. B. Hammer, Y. Morikawa, J. K. Nørskov, *Phys. Rev. Lett.* **1996**, *76*, 2141.
88. M. T. M. Koper, R. A. van Santen, S. A. Wasileski et al., *J. Chem. Phys.* **2000**, *113*, 4392.
89. M. T. M. Koper, in *Modern Aspects of Electrochemistry* (Eds.: C. G. Vayenas, B. E. Conway, R. E. White), Kluwer Academic/Plenum Publishers, New York, 2003, pp. 51–130, Vol. 36.
90. F. Illas, F. Mele, D. Curulla et al., *Electrochim. Acta* **1998**, *44*, 1213.
91. H. A. Gasteiger, N. M. Markovic, P. N. Ross Jr, *Catal. Lett.* **1998**, *36*, 617.
92. M. T. M. Koper, R. A. van Santen, *J. Electroanal. Chem.* **1999**, *476*, 64.
93. P. S. Bagus, G. Pacchioni, *Electrochim. Acta* **1991**, *36*, 1669.

94. M. T. M. Koper, T. J. Schmidt, N. M. Markovic et al., *J. Phys. Chem. B* **2001**, *105*, 8361.
95. S. K. Desai, P. Venkataraman, M. Neurock, *J. Phys. Chem.* **2001**, *105*, 7227.
96. M. Watanabe, New CO-tolerant catalyst concepts, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
97. M. Watanabe, H. Igarashi, T. Fujino, *Electrochemistry* **1999**, *67*, 1194.
98. H. Igarashi, T. Fujino, Y. Zhu et al., *Phys. Chem. Chem. Phys.* **2001**, *3*, 306.
99. C. G. Vayenas, S. Brosda, C. Pliangos, *J. Catal.* **2001**, *203*, 329–350.
100. S. Brosda, C. G. Vayenas, J. Wei, *Appl. Catal., B* **2006**, *68*, 109–124.
101. T. Iwasita, in *Advances in Electrochemical Science and Engineering* (Eds.: H. Gerischer, C. Tobias), Verlag Chemie, Weinheim, 1990, p. 127, Vol. 1.
102. M. Shibata, S. Motoo, *J. Electroanal. Chem.* **1985**, *192*, 367.
103. M. Watanabe, S. Motoo, *J. Electroanal. Chem.* **1975**, *60*, 259.
104. M. Watanabe, M. Horiuchi, S. Motoo, *J. Electroanal. Chem.* **1988**, *250*, 117.
105. J. O. M. Bockris, Electrode kinetics, in *Modern Aspects of Electrochemistry* (Ed.: J. O. M. Bockris), Scientific Publications, London, 1954, pp. 180–276, Chapter 4.
106. A. N. Frumkin, Hydrogen overvoltage and adsorption phenomena: part I, mercury, in *Advances in Electrochemistry and Electrochemical Engineering* (Eds.: P. Delahay, Ch. W. Tobias), Interscience Publishers, New York, 1961, pp. 65–122, Chapter 2.
107. A. N. Frumkin, Hydrogen overvoltage and adsorption phenomena: part II, in *Advances in Electrochemistry and Electrochemical Engineering* (Eds.: P. Delahay, C. W. Tobias), Interscience Publishers, New York, 1961, pp. 65–122, Chapter 2.
108. K. J. Vetter, *Électrochemische Kinetic*, Springer-Verlag, Berlin, 1961
109. T. Erdey-Gruz, M. Volmer, *Z. Phys. Chem. A* **1930**, *150*, 203.
110. J. Tafel, *Z. Phys. Chem.* **1905**, *50*, 641.
111. J. Heyrovsky, *Recl. Trav. Chem. Pays-Bas* **1927**, *46*, 582.
112. M. Temkin, *Zh. Fiz. Khim.* **1941**, *15*, 296.
113. A. N. Frumkin, *Z. Phys.* **1926**, *34*, 792.
114. M. W. Breiter, *Electrochemical Processes in Fuel Cells*, Springer-Verlag, Berlin, 1969.
115. M. W. Breiter, Reaction of the H_2 oxidation/evolution reaction in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
116. A. Lasia, Hydrogen evolution reaction in *Handbook of Fuel Cells*, Electrocatalysis, John Wiley and Sons, Chichester, 2003, Vol. 2.
117. N. M. Markovic, The hydrogen electrode reaction and the electrooxidation of CO and H_2/CO mixtures on well-characterized Pt and Pt-bimetallic surfaces, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 368–393, Vol. 2.
118. B. E. Conway, J. O. M. Bockris, *J. Chem. Phys.* **1957**, *26*, 532.
119. H. Gerischer, *Bull. Soc. Chem. Belg.* **1958**, *67*, 506.
120. S. Trasatti, *J. Electroanal. Chem.* **1972**, *39*, 163.
121. S. Trasatti, *Surf. Sci.* **1995**, *335*, 1.
122. J. Clavilier, A. Rhodes, M. A. Zamakhachari, *J. Chim. Phys.* **1991**, *88*, 1291.
123. M. Gattrell, B. MacDougall, Reaction mechanisms of the O_2 reduction/evolution reaction, in *Handbook of Fuel Cells*, Electrocatalysis, John Wiley and Sons, Chichester, 2003, Vol. 2.
124. P. N. Ross Jr, Oxygen reduction reaction on smooth single crystal electrodes, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
125. E. Ivers-Tiffée, A. Weber, H. Schichleim, Electrochemical impedance spectroscopy, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2.
126. J. P. Hoare, *The Electrochemistry of Oxygen*, Interscience, New York, 1968, 86.
127. A. Damjanovic, Mechanistic analysis of oxygen electrode reactions, in *Modern Aspects of Electrochemistry*, No. 5 (Eds.: B. E. Conway, J. O. M. Bockris), Plenum Press, New York, 1969, pp. 369–483, Chapter 5.
128. J. P. Hoare, Oxygen, in *The Encyclopedia of the Electrochemistry of the Elements* (Ed.: A. J. Bard), Dekker, New York, 1974, p. 208, Vol. 2.
129. M. R. Tarasevich, A. Sadkowski, E. Yeager, Oxygen electrochemistry, in *Comprehensive Treatise of Electrochemistry* (Eds.: B. E. Conway, J. O. M. Bockris, E. Yeager

et al.), Plenum Press, New York, 1983, pp. 301–398, Vol. 7.
130. A. J. Appleby, *J. Electroanal. Chem.* **1993**, *357*, 117.
131. A. Damjanovic, Progress in the studies of oxygen reduction during the last thirty years, in *Electrochemistry in Transition* (Eds.: O. J. Murphy, S. Srinivasan, B. E. Conway), Plenum Press, New York, 1992, pp. 107–126.
132. K. Kinoshita, *Electrochemical Oxygen Technology*, John Wiley and Sons, New York, 1992, pp. 19–65, 99–112.
133. N. M. Markovic, T. J. Schmidt, V. Stamenkovic et al., *Fuel Cells* **2001**, *1*, 105.
134. R. Adzic, Recent advances in the kinetics of oxygen reduction, in *Electrocatalysis* (Eds.: J. Lipkowski, P. N. Ross), Wiley-VCH, New York, 1998, pp. 197–242.
135. T. E. Springer, M. S. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1993**, *140*(12), 3512.
136. M. Pourbaix, *Atlas of Electrochemical Equilibria in Aqueous Solutions*, 2nd ed., NACE, Houston, 1974, pp. 97–105.
137. J. B. Goodenough, B. L. Cushing, Oxide-based ORR catalysts, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2, Chapter 35.
138. H. Angerstein-Kozlowska, B. E. Conway, W. B. A. Sharp, *J. Electroanal. Chem.* **1973**, *43*, 9.
139. S. Mukerjee, S. Srinivasan, M. P. Soriaga et al., *J. Electrochem. Soc.* **1995**, *142*(5), 1409.
140. T. Zambelli, J. V. Barth, J. Wintterlin et al., *Nature* **1997**, *390*(4), 495.
141. G. Chen, D. A. Delafuente, S. Sarangapani et al., *Catal. Today* **2001**, *67*, 341–355.
142. T. Iwasita, Methanol and CO electrooxidation, in *Handbook of Fuel Cells, Electrocatalysis*, John Wiley and Sons, Chichester, 2003, Vol. 2, Chapter 41.
143. M. Breiter, *Electrochim. Acta* **1967**, *12*, 1213.
144. T. Iwasita, in *Advances in Electrochemical Science and Engineering* (Eds.: H. Gerischer, C. Tobias), Verlag Chemie, Weinheim, 1990, p. 127, Vol. 1.
145. H. Baltruschat, N. A. Anastasijevic, M. Beltowska-Brzezinska et al., *Ber. Bunsen-Ges. Phys. Chem.* **1990**, *94*, 996–1000.
146. C. H. Hamann, A. Hamnett, W. Vielstich, Electrochemical modification of catalytic activity in heterogeneous chemical reactions-the NEMCA effect, in *Electrochemistry*, 2nd ed., Wiley-VCH, Weinheim, 1998, pp. 328–330.
147. W. Chrzanowski, A. Wieckowski, *Langmuir* **1998**, *14*, 1967–1970.
148. E. Leiva, C. Sánchez, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, pp. 93–131, Vol. 2.
149. E. Leiva, T. Iwasita, E. Herrero et al., *Langmuir* **1997**, *13*, 6287–6293.
150. S. Motoo, M. Watanabe, *J. Electroanal. Chem.* **1976**, *69*, 429–431.
151. J. Clavilier, R. Parsons, R. Durand et al., *J. Electroanal. Chem.* **1981**, *124*, 321.
152. C. Lamy, J. M. Leger, J. Clavilier et al., *J. Electroanal. Chem.* **1983**, *150*, 321.
153. R. R. Adjic, A. V. Tripkovic, W. O'Grady, *Nature* **1982**, *296*, 137.
154. R. R. Adzic, D. N. Simic, R. R. Despic et al., *J. Electroanal. Chem.* **1975**, *65*, 587.
155. E. Herrero, V. Climent, J. M. Feliu, *Electrochem. Commun.* **2000**, *2*, 636.
156. Y.-Y. Yang, Z. Y. Zhou, S.-G. Sun, *J. Electroanal. Chem.* **2001**, *500*, 233.
157. E. Leiva, T. Iwasita, E. Herrero et al., *Langmuir* **1997**, *13*, 6287.
158. A. Hamnett, in *Interfacial Electrochemistry: Theory, Experiment, and Applications* (Ed.: A. Wieckowski), Marcel Dekker, New York, 1999, p. 843.
159. W. Chrzanowski, A. Wieckowski, in *Interfacial Electrochemistry, Theory, Experiment, and Applications* (Ed.: A. Wieckowski), Marcel Dekker, New York, 1999, pp. 937–954.
160. C. Lamy, E. M. Belgsir, J. M. Leger, *J. Appl. Electrochem.* **2001**, *31*, 799.
161. P. Waszczuk, C. Coutanceau, J.-M. Leger et al., *J. Electroanal. Chem.* **2001**, *511*, 55.
162. N. M. Markovic, H. A. Gasteiger, P. N. Ross et al., *Electrochim. Acta* **1995**, *40*, 91.
163. N. M. Markovic, C. A. Lucas, V. Climent et al., *Surf. Sci.* **2000**, *465*, 103.
164. N. M. Markovic, V. Radmilovic, P. N. Ross, in *Catalysis and Electrocatalysis at Nanoparticle Surfaces* (Ed.: A. Wieckowski), Marcel Dekker, New York, 2002.
165. O. A. Petrii, Oxidation of C_2 molecules, in *Handbook of Fuel Cells, Electrocatalysis* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, Vol. 2, Chapter 45.
166. Y. Jiang, A. V. Virkar, *J. Electrochem. Soc.* **2001**, *148*, A706.

167. J. T. Muller, P. M. Urban, W. F. Hoelderich et al., *J. Electrochem. Soc.* **2000**, *147*, 4058.
168. A. Oliveira-Neto, M. J. Giz, J. Perez et al., *J. Electrochem. Soc.* **2002**, *149*, A272.
169. E. Peled, V. Livshits, T. Duvdevani, *J. Power Sources* **2002**, *106*, 245.
170. G. Foti, Ch. Comninellis, in *Modern Aspects of Electrochemistry* (Eds.: C. G. Vayenas, B. E. Conway, R. E. White), Kluwer Academic/Plenum Publishers, New York, 2003, pp. 3–130, Vol. 37.
171. C. Comninellis, A. De Battisti, *J. Chem. Phys.* **1996**, *93*, 673.
172. G. Foti, D. Gandini, C. Comninellis et al., *Electrochem. Solid-State Lett.* **1999**, *2*, 228.
173. F. Cardarelli, P. Taxil, A. Savall et al., *J. Appl. Electrochem.* **2001**, *148*, D78.
174. P.-A. Gherardini, M. Michaud, M. Panizza et al., *J. Electrochem. Soc.* **1998**, *281*, 245.
175. R. Chevrel, M. Sergent, J. Pringent, *J. Solid State Chem.* **1971**, *3*, 515.
176. (a) N. Alonso-Vante, H. Tributsch, *Nature (London)* **1986**, *323*, 431; (b) N. Alonso-Vante, W. Jaegermann, H. Tributsch et al., *J. Am. Chem. Soc.* **1987**, *109*, 3251; (c) N. Alonso-Vante, B. Schubert, H. Tributsch et al., *J. Catal.* **1988**, *112*, 384; (d) N. Alonso-Vante, B. Schubert, H. Tributsch, *Mater. Chem. Phys.* **1989**, *22*, 281; (e) C. Fischer, N. Alonso-Vante, S. Fiechter et al., *J. Appl. Electrochem.* **1995**, *25*, 1004.
177. N. Alonso-Vante, P. Bogdanoff, H. Tributsch, *J. Catal.* **2000**, *190*, 240.
178. R. Jasinski, *Nature* **1964**, *201*, 1212.
179. J. H. Zagal, *Coord. Chem. Rev.* **1992**, *119*, 89.
180. J. H. Zagal, G. Cardenas-Jiron, *Electrochim. Acta* **2000**, *489*, 96.
181. W. Vielstich, V. A. Paganin, F. H. B. Lima et al., *J. Electrochem. Soc.* **2001**, *148A*, 502.
182. J. T. Irvine, P. R. Slater, A. Kaiser et al., *Proceedings of the 4th European SOFC Forum Luzern*, Switzerland, 2000, p. 471.
183. (a) S. Park, R. Cracium, J. M. Vohs et al., *J. Electrochem. Soc.* **1999**, *146*, 3603; (b) C. Wang, W. L. Worrell, S. Park et al., *J. Electrochem. Soc.* **2001**, *146*, A864; (c) R. J. Gorte, H. Kim, J. M. Vohs, *J. Power Sources* **2002**, *106*, 10.
184. H. Kim, S. Park, J. M. Vohs et al., *J. Electrochem. Soc.* **2001**, *148*, A693.
185. E. Pastor, S. Wasmus, T. Iwasita et al., *J. Electroanal. Chem.* **1994**, *371*, 167.
186. E. Pastor, J. L. Rodriguez, Hydrogenation reactions, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications, Electrocatalysis* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, Vol. 2, Chapter 47.
187. Y. Hori, CO_2-reduction, catalyzed by metal electrodes, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications, Electrocatalysis* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, Chichester, 2003, Vol. 2, Chapter 48.
188. I. Taniguchi, Electrochemical and photoelectrochemical reduction of carbon dioxide, in *Modern Aspect of Electrochemistry*, Plenum Press, New York, 1989, p. 327, Vol. 20.
189. Y. Hori, A. Murata, R. Takahashi, *J. Chem. Soc., Faraday Trans. I* **1989**, *85*, 2309.
190. D. W. DeWulf, A. J. Bard, *Catal. Lett.* **1988**, *1*, 73.
191. S. Ikeda, *Bull. Chem. Soc. Jpn.* **1987**, *60*, 2517.
192. Y. Hori, H. Wakebe, T. Tsukamoto et al., *Electrochim. Acta* **1994**, *39*, 1833.
193. B. N. Nikolic, H. Huang, D. Gervasio et al., *J. Electroanal. Chem.* **1990**, *295*, 415.
194. K. Hara, A. Kudo, T. Sakata et al., *J. Electrochem. Soc.* **1995**, *142*, L57.
195. C. G. Vayenas, S. Bebelis, S. Ladas, *Nature* **1990**, *343*, 625–627.
196. C. G. Vayenas, M. M. Jaksic, S. Bebelis et al., *The Electrochemical Activation of Catalysis* (Eds.: J. O. M. Bockris, B. E. Conway, R. E. White), Kluwer Academic/Plenum Publishers, New York, 1996, No. 29, pp. 57–202.
197. J. Pritchard, *Nature* **1990**, *343*, 592.
198. R. M. Lambert, F. Williams, A. Palermo et al., *Top. Catal.* **2000**, *13*, 91.
199. G. Foti, I. Bolzonella, C. Comninellis, in *Modern Aspects of Electrochemistry* (Eds.: C. G. Vayenas, B. E. Conway, R. E. White), Kluwer Academic/Plenum Publishers, New York, 2003, pp. 191–254, Vol. 36.
200. C. A. Cavalca, G. L. Haller, *J. Catal.* **1998**, *177*, 389.
201. L. Ploense, M. Salazar, B. Gurau et al., *J. Am. Ceram. Soc.* **1997**, *119*, 11550.
202. P. Vernoux, F. Gaillard, L. Bultel et al., *J. Catal.* **2002**, *208*, 412.
203. I. Metcalfe, *J. Catal.* **2001**, *199*, 247.
204. G.-Q. Lu, A. Wieckowski, *Curr. Opin. Colloid Interf. Sci.* **2000**, *5*, 95–100.

205. B. Grzybowska-Swierkosz, J. Haber, *Annual Reports on the Progress of Chemistry*, The Royal Society of Chemistry, Cambridge, 1994.
206. D. Tsiplakides, S. Balomenou, A. Katsaounis et al., *Catal. Today* **2005**, *100*, 133–144.
207. C. G. Vayenas, S. Brosda, C. Pliangos, 40th Anniversary issue, *J. Catal.* **2003**, *216*, 487–504.
208. C. Sánchez, E. Leiva, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, New York, 2003, pp. 65–68, Vol. 2.
209. E. Leiva, C. Sánchez, in *Handbook of Fuel Cells–Fundamentals, Technology and Applications* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), John Wiley and Sons, New York, 2003, pp. 145–149, Vol. 2.
210. S. Bebelis, C. G. Vayenas, *J. Catal.* **1989**, *118*, 125–146.
211. C. Kokkofitis, G. Karagiannakis, M. Stoukides, *Top. Catal.* **2006**, in press.
212. N. Kotsionopoulos, S. Bebelis, *Top. Catal.* **2006**, in press.
213. D. L. Rath, D. M. Kolb, *Surf. Sci.* **1981**, *109*, 641–647.
214. D. M. Kolb, *Z. f. Phys. Chem. Neue Folge* **1987**, *154*, 179–199.
215. I. Riess, C. G. Vayenas, *Solid State Ionics* **2003**, *159*, 313–329.
216. P. K. Babu, E. Oldfield, A. Wieckowski, in *Modern Aspects of Electrochemistry* (Eds.: C. G. Vayenas, B. E. Conway, R. E. White), Kluwer Academic/Plenum Publishers, New York, 2003, pp. 1–50, Vol. 36.

3
Electrochemical Deposition

Milan Paunovic
IBM Research Division Yorktown Heights, New York

3.1	**Electrodeposition**	**91**
3.1.1	Introduction	91
3.1.2	Metal/Metal-ion (M/M^{z+}) Interphase at Equilibrium	92
3.1.3	Electrode Potential Departs from Equilibrium: Overpotential	93
3.1.3.1	High-overpotential Approximation	95
3.1.3.1.1	Large Cathodic Current	95
3.1.3.1.2	Large Anodic Current	95
3.1.3.2	Low-overpotential Approximation	97
3.1.3.3	Influence of Mass Transport on Electrode Kinetics	97
3.1.3.4	Underpotential Deposition	98
3.1.4	Atomistic Aspects of Electrodeposition	99
3.1.4.1	Step-edge Ion-transfer Mechanism	99
3.1.4.2	Terrace Ion-transfer Mechanism	99
3.1.5	Techniques for the Study on Electrode Processes	101
3.1.5.1	Galvanostatic Transient Technique	101
3.1.5.2	Potentiostatic Transient Technique	102
3.2	**Nucleation and Growth**	**103**
3.2.1	Introduction	103
3.2.2	Nucleation and Growth of Surface Nuclei	103
3.2.2.1	Nucleation of Surface Nuclei	104
3.2.2.2	Growth of Surface Nuclei	105
3.2.3	Layer Growth Mechanism	106
3.2.3.1	Monoatomic Steps	106
3.2.3.2	Microsteps	106
3.2.3.3	Macrosteps	107
3.2.3.4	Lateral Merging of Steps	107
3.2.4	Nucleation-coalescence Growth Mechanism	108
3.2.4.1	Coalescence-induced Defects	108
3.2.4.2	Deposition of a Metal on a Foreign Metallic Substrate	108

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

3.2.5	Effect of Additives	109
3.2.5.1	Chemisorption and Physisorption	109
3.2.5.2	Adsorption Equilibrium	110
3.2.5.3	Effect of Additives on Kinetics and Mechanism of Electrodeposition	110
3.2.5.4	Effect of Additives on Nucleation and Growth	110
3.2.5.5	Dependence of Types of Deposit on the Surface Coverage by Additive	111
3.2.5.6	Leveling	111
3.2.5.7	Brightening	111
3.2.5.8	Consumption of Additives	111
3.3	**Electroless Deposition**	**112**
3.3.1	Introduction	112
3.3.2	Electrochemical Model: the Mixed-potential Theory	113
3.3.2.1	Wagner–Traud Diagram	114
3.3.2.2	Evans Diagram	115
3.3.2.3	Mixed Potential, E_{mp}	115
3.3.3	Kinetics and Mechanism	117
3.3.3.1	The Cathodic Partial Reaction. Kinetic Scheme	117
3.3.3.2	Kinetics	117
3.3.3.3	The Anodic Partial Reaction. Mechanism	117
3.3.3.4	Parallel Reactions	118
3.3.3.5	Kinetics	118
3.3.4	Activation of Noncatalytic Surfaces	119
3.3.4.1	Electrochemical Activation	119
3.3.4.2	Photochemical Activation	119
3.3.4.3	Activation by Displacement Deposition	119
3.3.4.4	Activation by Thermal Decomposition of Metallic Oxides	119
3.4	**Displacement Deposition**	**120**
3.4.1	Introduction	120
3.4.2	Electrochemical Model	121
3.4.2.1	Predictions of the Feasibility of the Reaction. Complexed Metal Ions in Displacement Deposition	122
3.4.3	Kinetics and Mechanism	123
3.5	**Electrochemical Deposition of Alloys**	**123**
3.5.1	Introduction	123
3.5.2	Electrodeposition	124
3.5.2.1	Feasibility	124
3.5.2.2	A Simple Model	124
3.5.2.3	Extended Feasibility	125
3.5.2.4	Anomalous Codeposition	126
3.5.3	Electroless Deposition	128

3.6	**Structure and Properties of Deposits**	**129**
3.6.1	Introduction	129
3.6.2	Overpotential Dependence of the Growth Forms	129
3.6.3	Columnar Microstructure	129
3.6.4	Development of Texture	131
3.6.4.1	Single-crystal Substrate	131
3.6.4.2	Polycrystalline Randomly Oriented Substrate	131
3.6.4.3	Amorphous Substrate	132
3.6.5	Dendrites	132
3.6.6	Whiskers	132
3.6.7	Electrical Resistance	133
3.6.8	Mechanical Properties	134
3.7	**Applications in Semiconductor Technology**	**134**
3.7.1	Introduction	134
3.7.2	Deposition of Cu Interconnections on Chips	136
3.7.3	Diffusion Barriers and Seed Layer	138
3.8	**Applications in Magnetic Recording Industry**	**142**
3.8.1	Introduction	142
3.8.2	Write Heads	143
3.8.3	Read Heads	145
3.8.4	Magnetic Recording Media	149
	References	152
	Appendix	156
A1	**Magnetization and Hysteresis Loops for Hard Magnetic Material**	**156**
A2	**Magnetization and Hysteresis Loops for Soft Magnetic Material**	**157**
A3	**Magnetic Anisotropy**	**157**
A4	**Magnetostriction**	**158**

Electrochemical deposition involves (1) electrodeposition (Sects 3.1, 3.2, and 3.5.2), (2) electroless deposition (Sects 3.3 and 3.5.3), and (3) displacement deposition (Sect. 3.4). The structure and properties of the deposits are the subject of Sect. 3.6. The applications of electrochemical deposition in modern electrochemical technology are discussed in Sects 3.7 and 3.8.

3.1 Electrodeposition

3.1.1 Introduction

Electrodeposition of metals involves the reduction of metal ions from aqueous, organic, and fused-salt electrolytes. The reduction of metal ions M^{z+} in aqueous solution is represented by

$$M^{z+}_{\text{solution}} + ze \longrightarrow M_{\text{lattice}} \qquad (1)$$

where z is the number of electrons (e).

A simplified representation of the electrodeposition process, Eq. (1), is

$$M^{z+}_{\text{solution}} \longrightarrow M^{z+}_{\text{lattice}} \qquad (2)$$

since a metal may be considered a fixed lattice of positive ions permeated by a gas of free electrons [1]. According to this representation, a metal ion M^{z+} is transferred from the solution into the ionic metal lattice. This process is accompanied by the transfer of z electrons (e) from an external electron source (power supply) to the electron gas in the metal M [1, 2].

The overall reaction, Eq. (1), may take place in a number of steps or partial reactions. There are four possible partial reactions: charge transfer, mass transport, chemical reaction, and crystallization. Charge-transfer reactions involve the transfer of charge carriers (ions or electrons) across the double layer. This is the basic deposition reaction. The charge-transfer reaction is the only partial reaction directly affected by the electrode potential. In mass transport processes, the substances consumed or formed during the electrode reaction are transported from the bulk solution to the interphase (double layer) and from the interphase to the bulk solution. This mass transport takes place by diffusion. Chemical reactions involved in the overall deposition process can be homogeneous reactions in the solution and heterogeneous reactions at the surface. The rate constants of chemical reactions are independent of the potential. In crystallization partial reactions, atoms are either incorporated into or removed from the crystal lattice.

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

The slowest partial reaction is rate determining for the total overall reaction.

Section 3.1.2 discusses metal/metal-ion interphase at equilibrium; Sect. 3.1.3 discusses the relationship between current and potential; Sect. 3.1.4 describes the atomistic aspects of electrodeposition and Sect. 3.1.5 presents the techniques for studying deposition processes.

3.1.2 Metal/Metal-ion (M/M^{z+}) Interphase at Equilibrium

Here we consider a metal M immersed in the aqueous solution of its salt MA. In this case, both the phases, metal and the ionic solution, contain M^{z+} ions. Thus, there will be an exchange of metal ions M^{z+} between the two phases at the metal–solution interface (physical boundary) (Fig. 1).

Let us assume that conditions are such that more M^{z+} ions leave than enter the crystal lattice. In this case, there is an excess of electrons on the metal and the metal acquires negative charge, $q_M{}^-$ (charge on the metal per unit area). In response to the charging of the metal side of the interface, there is a rearrangement of charges on the solution side of the interface. The negative charge on the metal attracts positively charged M^{z+} ions from the solution and repels negatively charged A^{z-} ions. The result of this is an excess of positive M^{z+} ions in the solution in the vicinity of the metal interface. If the number per square centimeter of ionic species i in the bulk of solution is n_i^b and the number per square centimeter of these species in the interphase is n_i, then the excess charge of ionic species in the interphase is

$$\Delta n_i = n_i - n_i^b \tag{3}$$

Thus, in this case the solution side of the interphase acquires an opposite and equal charge, $q_S{}^+$ (the charge per unit area on the solution side of the interphase). At *the dynamic equilibrium*, the interphase region is neutral

$$q_M = -q_S \tag{4}$$

and the same number of M^{z+} ions enter, $\overleftarrow{n}$, and the same number of M^{z+} ions leave the crystal lattice, $\overrightarrow{n}$, (Fig. 1)

$$\overrightarrow{n} = \overleftarrow{n} \tag{5}$$

In the general case of a metal/metal-ion electrode, a metal M is in equilibrium with its ions in the solution

$$M^{z+} + ze \longleftrightarrow M \tag{6}$$

The reaction from left to right consumes electrons and is called *reduction*. The reaction from right to left liberates electrons and is called *oxidation*.

The next question is how the excess charges q_M and $-q_S$ are distributed on the metal and solution sides of the interphase.

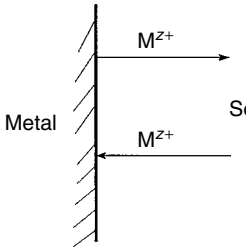


Fig. 1 Metal/metal-ion interphase at equilibrium: an equal number of metal ions M^{z+} crossing in both directions across the metal/solution interphase.

These topics are discussed in Chapter 3, Volume 1. The potential difference across the interphase is $\Delta\Phi(\mathrm{M, S}) = \Phi_\mathrm{M} - \Phi_\mathrm{S}$. This potential difference cannot be measured directly (see Chapter 1 in Volume 1). The electrode potential E of the metal/metal-ion electrode (with respect to the reference electrode) depends on the concentration (more exactly, the activity) of the metal ions M^{z+} in solution according to the Nernst equation

$$E = E^\mathrm{o} + \frac{RT}{zF} \ln[\mathrm{M}^{z+}] \qquad (7)$$

where R, T, z, and F are the gas constant, absolute temperature, number of electrons involved in reaction (6), and Faraday's constant (96 500 C), respectively. The square brackets [], in general, signify activity of the species inside the brackets; when the concentration of the solution is low, for example, 0.001 M or lower, the activity in [] can be replaced by the concentration in moles per liter (the relationship between concentration and activity is discussed in Volume 1). When the activity of M^{z+} in the solution is equal to 1, then from Eq. (7) it follows that (since $\ln 1 = 0$)

$$E = E^\mathrm{o}, [\mathrm{M}^{z+}] = 1 \qquad (8)$$

where E^o is the relative standard electrode potential of the $\mathrm{M/M}^{z+}$ electrode.

Converting the natural logarithm into decimal logarithm, Eq. (7) yields

$$E = E^\mathrm{o} + 2.303\frac{RT}{zF} \log[\mathrm{M}^{z+}] \qquad (9)$$

The quantity RT/F has the dimensions of voltage and at 298 K (25 °C) has the value $(RT/F) = 0.0257$ V and the quantity 2.303 $(RT/F) = 0.0592$ V. With these values, Eq. (9) becomes

$$E = E^\mathrm{o} + \frac{0.0592}{z} \log[\mathrm{M}^{z+}] \qquad (10)$$

E^o can be determined by plotting $\log[\mathrm{M}^{z+}]$ against E and extrapolating to $[\mathrm{M}^{z+}] = 1$.

3.1.3 Electrode Potential Departs from Equilibrium: Overpotential

When an electrode is made a part of an electrochemical cell through which current is flowing, its potential will differ from the equilibrium potential. If the equilibrium potential of the electrode (potential in the absence of external current) is E and the potential of the same electrode as a result of external current I flowing is $E(I)$, then the difference η between these two potentials

$$\eta = E(I) - E \qquad (11a)$$

is called *overpotential*. In terms of $\Delta\Phi$

$$\eta = \Delta\Phi(I) - \Delta\Phi_\mathrm{eq.} \qquad (11b)$$

The overpotential η is required to overcome the hindrance of the overall electrode reaction, which is usually composed of the sequence of partial reactions. There are four possible partial reactions, as described in Sect. 3.1.1 and, thus, four types of rate control: charge transfer, diffusion, chemical reaction, and crystallization. Four different kinds of overpotential are distinguished and the total overpotential η can be considered to be composed of four components

$$\eta = \eta_\mathrm{ct} + \eta_\mathrm{d} + \eta_\mathrm{r} + \eta_\mathrm{c} \qquad (12)$$

where η_ct, η_d, η_r, and η_c are, as defined above, charge-transfer, diffusion, chemical reaction, and crystallization overpotentials, respectively.

We first consider the case in which the charge transfer is the slow step. In this case, the rate of the electrode reaction is determined by the charge-transfer overpotential, $\eta = \eta_\mathrm{ct}$. It is assumed here that

other processes, such as mass transport, are fast steps.

Work on the development of the modern theory of the charge-transfer overpotential started when Eyring and Wynne-Jones and Eyring formulated the absolute rate theory on the basis of statistical mechanics [3, 4]. This expresses the rate constant k of a chemical reaction in terms of the activation energy $\Delta G^{\neq}$, Boltzmann's constant k_B and Planck's constant h

$$k = \frac{k_B T}{h} \exp(-\Delta G^{\neq}/RT) \qquad (13)$$

where R and T are the gas constant and absolute temperature, respectively.

For an electrochemical reaction, the rate and the rate constant depend on electrode potential E, specifically, the potential difference across electrode–solution interphase $\Delta\Phi$ through *the electrochemical activation energy* $\Delta G^{\neq}_{ec}$. Thus, the central problem here is to find the function

$$\Delta G^{\neq}_{ec} = f(E) \qquad (14)$$

It is not possible at present to evaluate this function theoretically. The system is too complicated. An approximate solution is possible if we assume that $\Delta G^{\neq}_{ec}$ can be separated into two parts

$$\Delta G^{\neq}_{ec} = \Delta G^{\neq}_{pin} + \Delta G^{\neq}_{pd} \qquad (15)$$

where $\Delta G^{\neq}_{pin}$ and $\Delta G^{\neq}_{pd}$ are the potential-independent and the potential-dependent parts of the electrochemical activation energy, respectively. The potential-dependent part of the activation energy $\Delta G^{\neq}_{pd}$ can be estimated (approximated) by the introduction of the transfer coefficient α, which was introduced in 1930 by Erdey-Gruz and Volmer [5]. With this approximation, the potential-dependent contribution for the reduction reaction is $\alpha z F \Delta\Phi$ where α is in the range from 0 to 1 and is often 0.5 (symmetric energy barrier). For the reverse reaction, oxidation, the potential-dependent contribution is $(1-\alpha) z F \Delta\Phi$. The substitution of these values into Eq. (15) yields $\Delta G^{\neq}_{ec} = f(E)$ for the reduction and the oxidation processes. Using these values for the electrochemical activation energy and equations for expressing the rate of electrochemical reaction in terms of the current density i (in amperes per square centimeter), one gets the *current–potential relationship for the partial reactions* [1]. Thus, the cathodic partial current density is

$$\vec{i} = i_o \exp(-\alpha z F \eta/RT) \qquad (16)$$

and the anodic partial current density is

$$\overleftarrow{i} = i_o \exp((1-\alpha) z F \eta/RT) \qquad (17)$$

where i_o is the exchange current density, one of the most important parameters of electrochemical kinetics.

When an electrode is at equilibrium, the equilibrium partial current densities $\vec{i}$ and $\overleftarrow{i}$ are equal and they are designated by one symbol, i_o. This equality on an atomic scale means that a constant exchange of charge carriers (electrons or ions) takes place across the metal–solution interphase (Fig. 1). When the interphase is not in equilibrium, a net current density i flows through the electrode (the double layer). It is given by the difference between the anodic partial current density $\overleftarrow{i}$ (a positive quantity) and the cathodic partial current density $\vec{i}$ (a negative quantity):

$$i = \overleftarrow{i} - \vec{i} \qquad (18)$$

Figure 2 illustrates the metal/metal-ion electrode in nonequilibrium state. From Eqs (16–18), we get the Butler–Volmer

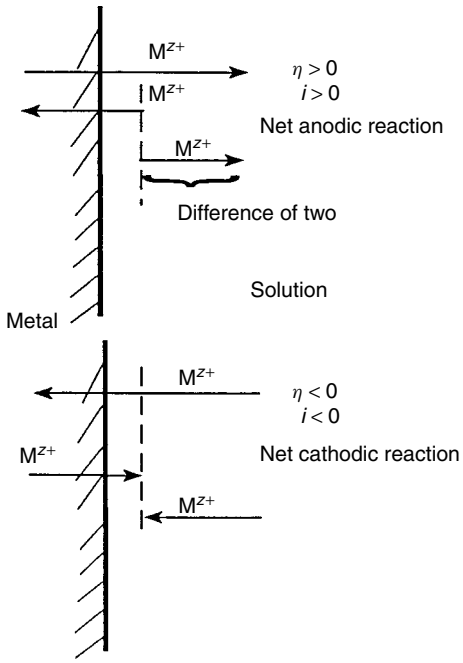


Fig. 2 Metal/metal-ion electrode deviates from the equilibrium potential: two currents flowing in opposite directions.

equation

$$i = i_o[\exp((1-\alpha)zF\eta/RT) - \exp(-\alpha zF\eta)/RT] \quad (19)$$

This equation gives the relationship between the current density i and the charge-transfer overpotential η in terms of two parameters, the exchange current density i_o and the transfer coefficient α. Figure 3 shows the variation of the partial current densities and the net current density with overpotential. It can be seen that for large departures from the equilibrium potential, large $\pm\eta$ values, the partial current densities approach the net current density i. For large positive values of overpotential, $\overleftarrow{i} \approx i$; and for large negative values of overpotential, $\overrightarrow{i} \approx i$. The double exponential Butler–Volmer equation has two limiting forms.

3.1.3.1 High-overpotential Approximation

3.1.3.1.1 **Large Cathodic Current** For these conditions Eq. (19) can be simplified. Analysis of Eq. (19) shows that when η becomes more negative, the first exponential term in the equation (corresponding to the anodic partial current) decreases, while the second exponential term (corresponding to the cathodic partial reaction) increases. Under these conditions, the smaller term $\exp[(1-\alpha)zF\eta/RT]$ can be neglected, yielding an approximate form of the Butler–Volmer equation which is essentially single exponential Eq. (16).

3.1.3.1.2 **Large Anodic Current** When η has large positive values, the second exponential term in Eq. (19) decreases, while the first term (corresponding to the anodic partial reaction) increases. Under these conditions, the smaller term $\exp[-\alpha zF\eta/RT]$ can be neglected, yielding an approximate form which is essentially single exponential Eq. (17).

Thus, there is an exponential relationship between i and high values of η, usually greater than 0.10 V. Taking the logarithm of both sides of Eqs (16 and 17) and expressing η as a function of the current density, we have

$$\eta = a \pm b \log|i| \quad (20)$$

where the $\pm$ sign holds for anodic and cathodic processes, respectively. It should be noticed that the theoretical values of the constants a and b are different for the two processes [1]. The linear relationship

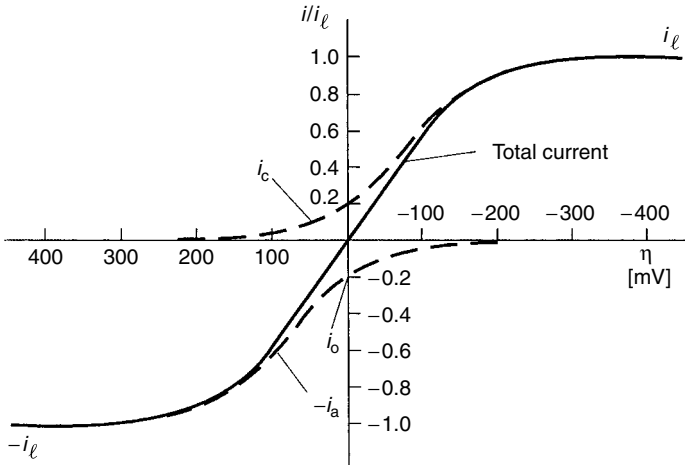


Fig. 3 The variation of partial current densities (- - - -) and the net current density (———) with overpotential η.

between η and log i was experimentally established in 1905 by Tafel [6].

The experimentally determined cathodic and anodic Tafel lines for the electrodeposition of copper in acid copper sulfate solution are given in Fig. 4. Data presented in Fig. 4 can be used to evaluate the Tafel slope b (Eq. (20) for the deposition and dissolution of copper. From the cathodic polarization curve, we obtain the Tafel slope $\vec{b}$ as

$$\vec{b} = [\partial\eta/\partial \log |i|] = [\Delta\eta/\Delta \log |i|] = 125 \quad (21)$$

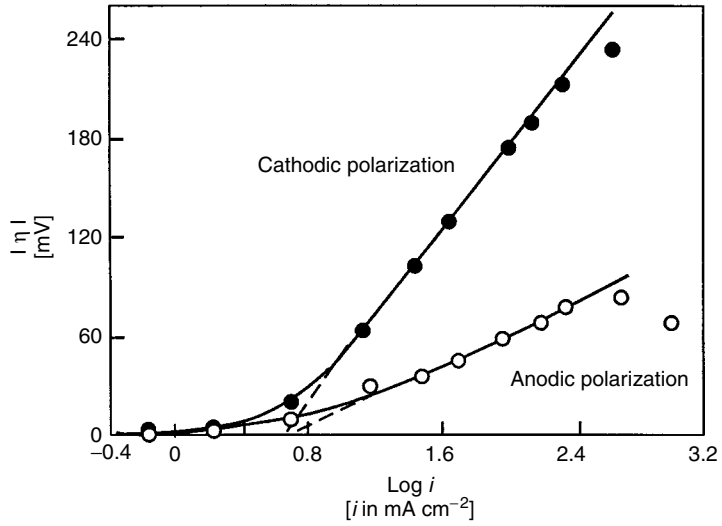


Fig. 4 Linear relationship between i and η for the electrodeposition of Cu from $CuSO_4-H_2SO_4$ solution (from Ref. 7 with permission from the Electrochemical Society).

and for the anodic polarization curve

$$\overleftarrow{b} = [\partial\eta/\partial \log|i|] = [\Delta\eta/\Delta \log|i|] = 50 \quad (22)$$

3.1.3.2 Low-overpotential Approximation

Under these conditions, the exponentials in the Butler–Volmer equation (19) can be approximated using a power series, and taking only the first-order terms in η, we get

$$i = i_o(zF/RT)\eta \quad (23)$$

or

$$i = i_o f \eta \quad (24)$$

where $f = F/RT$.

Thus, for small values of η, less than about 0.01 V, when the electrode potential is near the reversible potential, the current varies linearly with the overpotential.

3.1.3.3 Influence of Mass Transport on Electrode Kinetics

The current–potential relationships defined by Eqs (16, 17, and 19) are valid for the case in which the charge transfer is the slow process (rate-determining step, RDS). This relationship has a limit where the rate of deposition reaction is limited by the transport of M^{z+} ions. A general current–potential relationship is shown in Fig. 5. *The limiting diffusion current density* i_L, or maximum current density, is given by

$$i_L = (nFD/\delta)c_b \quad (25)$$

where D is the diffusion coefficient of the depositing species M^{z+}, c_b is the bulk concentration of M^{z+} ions in the solution, δ is the diffusion-layer thickness, n is the number of electrons involved in the reaction, and F is the Faraday's constant. The diffusion-layer thickness δ is defined by the Nernst diffusion layer model illustrated in Fig. 6. This model assumes that the concentration of M^{z+} ions has a bulk concentration c_b up to a distance δ from the electrode surface and then falls off linearly to $c_{x=0}$ at the electrode surface. In this model, it is assumed that the liquid layer of thickness δ is practically stationary (quiescent). At a distance greater than δ

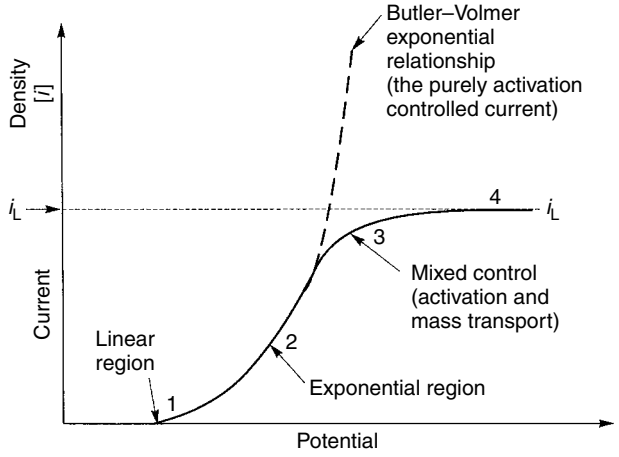


Fig. 5 Four regions in the general current–overpotential relationship: 1 – linear, 2 – exponential, 3 – mixed control, and 4 – limiting current density region.

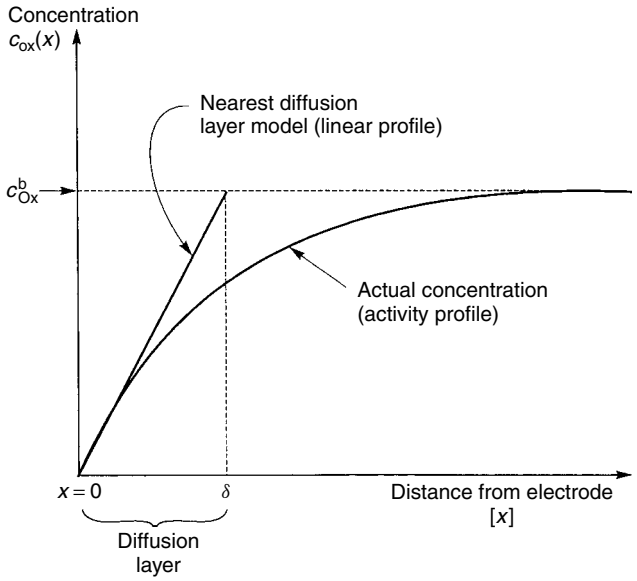


Fig. 6 Variation of concentration of the reactant during nonsteady-state electrolysis; c^b_{Ox} is the concentration in the bulk; $c_{Ox}(x)$ is the concentration at the surface.

from the surface, the concentration of the reactant M^{z+} is assumed to be equal to that in the bulk. At this distance, $x > \delta$, stirring is efficient. Ions M^{z+} must diffuse through the diffusion layer to reach the electrode surface.

At the value of the limiting (maximum) current density, the species M^{z+} are reduced as soon as they reach the electrode. At these conditions, the concentration of the reactant M^{z+} at the electrode is nil and the rate of deposition reaction is controlled by the rate of transport of the reactant M^{z+} to the electrode. If an external current greater that the limiting current i_L is forced through the electrode, the double layer is further charged, and the potential of the electrode will change until some process other than the reduction of M^{z+} can occur.

3.1.3.4 Underpotential Deposition

We have seen earlier that metal M will be deposited on the cathode from the solution of M^{n+} ions if the electrode potential E is more negative than the Nernst potential of the electrode M/M^{n+}. However, it is known that in many cases metal M can be deposited on a foreign substrate S from a solution of M^{n+} ions at potentials more positive than the Nernst potential of M/M^{n+}. This electrodeposition of metals is termed *underpotential deposition* (UPD). Thus, in terms of the actual electrode potential E during deposition and the Nernst equilibrium potential $E(M/M^{n+})$ and their difference $\Delta E = E - E(M/M^{n+})$, we distinguish two types of electrodeposition

1. Overpotential deposition (OPD)

$$E < E(M/M^{n+}),\ \Delta E < 0 \qquad (26)$$

2. Underpotential deposition (UPD)

$$E > E(M/M^{n+}),\ \Delta E > 0 \qquad (27)$$

UPD is discussed in detail in Chapter 6 in Volume 1.

3.1.4 Atomistic Aspects of Electrodeposition

Atomistic processes that constitute the electrodeposition process, Eqs (1 and 2), can be seen by presenting the structure of the initial, M^{z+} (solution), and the final, M^{z+} (lattice), states. Since metal ions in the aqueous solution are hydrated, the structure of the initial state in Eq. (2) is represented by $[M(H_2O)_x]^{z+}$. The structure of the final state is the M adion (atom) at the kink site (Fig. 7), since it is generally assumed that atoms (ions) are attached to the crystal via a kink site [1] (see Chapter 4.1 in Volume 2). Thus, the final step of the overall reaction, Eq. (2), is the incorporation of M^{z+} adions into the kink site. Because of surface inhomogeneity, the transition from the initial $[M(H_2O)_x]^{z+}$ (solution) to the final M^{z+} (kink) state

$$[M(H_2O)_x]^{z+}(\text{solution}) \longrightarrow M^{z+}(\text{kink}) \tag{28}$$

can proceed via either (1) step-edge site ion-transfer mechanism or (2) terrace site ion-transfer mechanism.

3.1.4.1 Step-edge Ion-transfer Mechanism

The step-edge site ion transfer, or direct transfer mechanism, is illustrated in Fig. 8. It shows that in this mechanism ion transfer from solution (OHP, outer Helmholtz plane of the double layer; see Chapter 3 in Volume 1) takes place on a kink site of a step edge or on any other site on the step edge. In both the cases, the result of the ion transfer is an M adion in the metal crystal lattice. In the first case, a direct transfer to the kink site, the M adion is in the half-crystal position, where it is bonded to the crystal lattice with one-half of the bonding energy of the bulk ion. Thus, the M adion belongs to the bulk crystal. However, it still has some water of hydration (Fig. 8). In the second case, a direct transfer to the step-edge site other than kink, the transferred metal ion diffuses along the step edge until it finds a kink site (Fig. 8).

Thus, in a step-edge site transfer mechanism, there are two possible paths: direct transfer to a kink site and the step-edge diffusion path.

3.1.4.2 Terrace Ion-transfer Mechanism

In the terrace site transfer mechanism, a metal ion is transferred from the solution

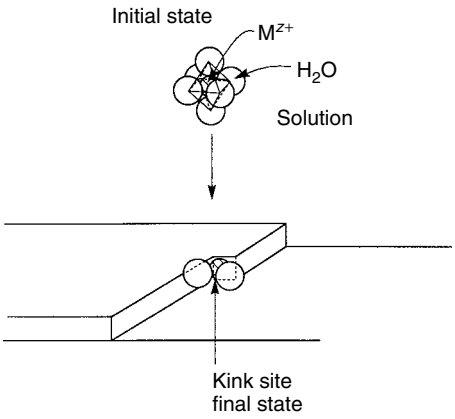


Fig. 7 Initial and final states in metal deposition.

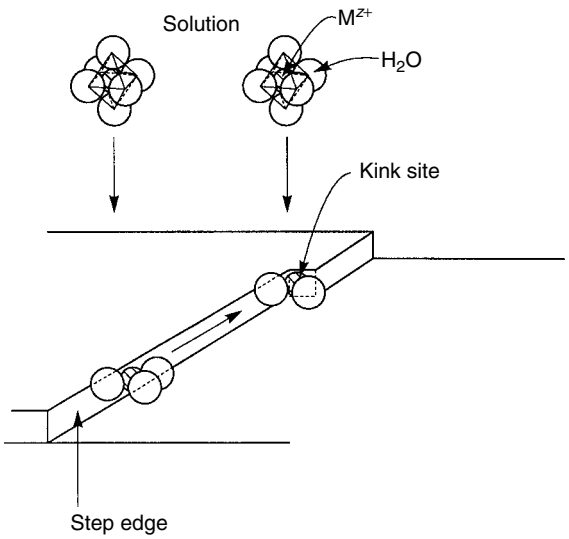


Fig. 8 Step-edge ion-transfer mechanism.

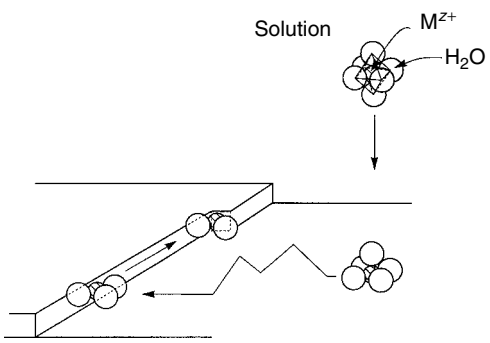


Fig. 9 Ion transfer to the terrace site, surface diffusion, and incorporation at kink site.

(OHP) to the flat surface of the terrace region (Fig. 9). At this position, the metal ion is in the adion (adsorbed-like) state having most of its water of hydration. It is weakly bound to the crystal lattice. From this position, it diffuses on the surface and loses some water of hydration, seeking a position of lower energy. The final position is a kink site.

In view of these two mechanisms, the step-edge and terrace ion transfer, it is clear that the basic factors in determining the kinetics and mechanism of the deposition process and the morphology of deposit are the atomic structure of the surface, the population density of the surface defects (dislocations, growth edges, kinks, and vacancies) and adsorbed species, and the electronic structure of surfaces (Ref. 1; Chapter 4 in Volume 2).

The initial theoretical treatment of these mechanisms of deposition was given by Lorenz [6, 8–10]. The initial experimental studies on surface diffusion were published by Mehl and Bockris [7, 11]. Conway and Bockris [12, 13] calculated activation energies for the ion-transfer process of various surface sites. The simulation of

crystal growth with surface diffusion was discussed by Gilmer and Bennema [14].

3.1.5 Techniques for the Study on Electrode Processes

The study of the charge-transfer processes, free from the effects of mass transport, is possible by the use of transient techniques. In the simplest transient techniques, the interface at equilibrium is abruptly changed from an equilibrium state to a steady state characterized by a new potential difference $\Delta\Phi$. This abrupt change is imposed by applying either (1) the constant current, galvanostatic transient technique or (2) the constant potential, potentiostatic transient technique. The study of transient processes utilizes the three-electrode cell design (Fig. 10).

3.1.5.1 Galvanostatic Transient Technique

In the galvanostatic technique, the current between the test electrode and the auxiliary (counter) electrode is held constant with a current source (galvanostat) and the potential between the test electrode and the reference electrode is determined as a function of time. The potential is the dependent variable, which is recorded with a suitable recording system such as recorders or oscilloscopes.

The input signal and the response to the signal are compared in Fig. 11. The response function $E_i = f(t)$ shows that a certain time (t_1) is necessary to reach a potential E_i when the electrode reaction begins at a measurable rate. When a constant current is applied, the current (electron flow) is used for (1) charging the double-layer capacitance C_{dl} up to the potential at which the electrode reaction can proceed with a measurable velocity and (2) electrode reaction (charge transfer). Thus, the total galvanostatic current density i_g is given by

$$i = i_{dl} + i_{ct} \tag{29}$$

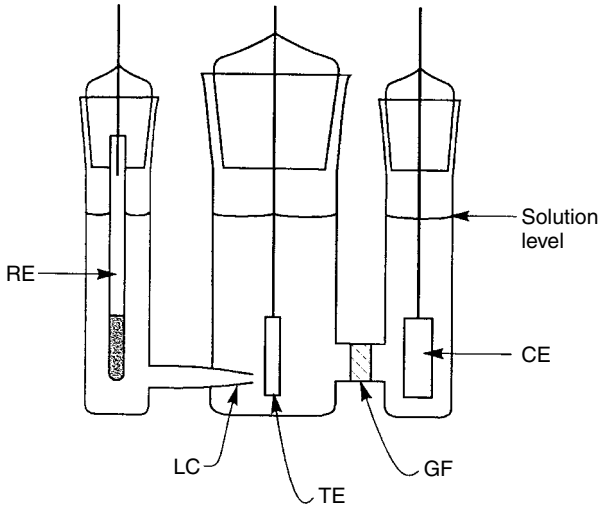


Fig. 10 Three-compartment, three-electrode, electrochemical cell; RE, reference electrode; LC, Lugin capillary; TE, test electrode; GF, glass frit; CE, counterelectrode.

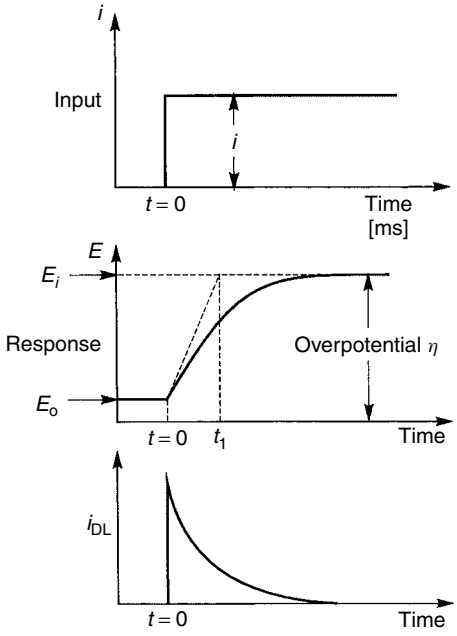


Fig. 11 Variation of the potential of the test electrode, E, with time during galvanostatic electrolysis (millisecond range); E_o, equilibrium potential; E_i, potential of the test electrode at the beginning of electrolysis at constant current density i.

where i_{dl} is the capacitive and i_{ct} is the faradaic current (charge transfer).

The first process after applying a current is charging the double-layer capacitance, C_{dl}, from the reversible potential E up to the potential E_i when the electrode reaction begins at a measurable rate. The time necessary to charge capacitor C_{dl} is in the range of microseconds. Thus, in the galvanostatic transient technique the duration of the input current density pulse is of the order of milliseconds (ms). From a series of measurements of E_i for a set of i values, one can construct the current–potential relationship for an electrochemical process. Figure 4 was constructed in this way.

3.1.5.2 Potentiostatic Transient Technique

In the potentiostatic technique, the potential of the test electrode is controlled, while the current, the dependent variable, is measured as a function of time. The potential difference between the test electrode and the reference electrode is controlled by a potentiostat.

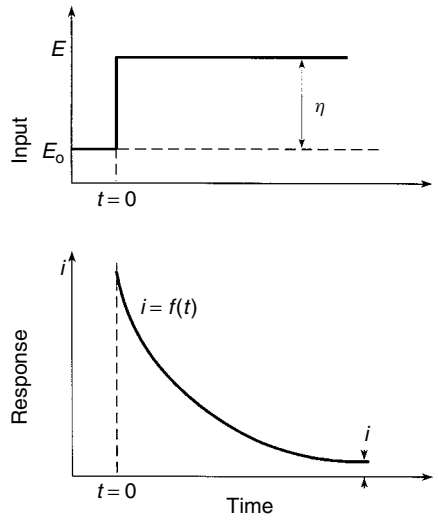


Fig. 12 Variation of current with time during potentiostatic electrolysis.

The input function, a constant potential, and the response function, $i = f(t)$, are shown in Fig. 12.

Other techniques for the study of electrode processes include, for example, potential sweep, alternating current (ac), and rotating electrodes [15].

3.2 Nucleation and Growth

3.2.1 Introduction

In Sect. 3.1.4, we presented atomistic processes in the transition of an *individual metal ion* from the solution to a position in a metal crystal lattice.

In this section, we discuss the building of a stable metal crystal lattice from *a sequence of a number of individual atomistic steps*. The building of a metal crystal lattice involves two groups of processes: (1) nucleation and growth of surface nuclei and (2) formation of coherent deposit. The first group includes the formation and growth of adion clusters and nucleation of surface nuclei. The formation of a coherent deposit may proceed by two basic mechanisms: layer growth and nucleation-coalescence growth. All these processes are influenced by the presence of additives in the solution.

Theoretical and experimental studies of the initial stages of building of a metal crystal lattice introduced a major advancement in the interpretation of the deposition processes and resulted in a series of new models [1, 16]. Experimental studies include applications of in situ surface analytical methods [17–21], including scanning tunneling microscopy (STM) and atomic force microscopy (AFM).

Initial stages of electrocrystallization are discussed in Sect. 3.2.2, growth mechanisms are discussed in Sects 3.2.3 and 3.2.4, and the effect of additives on these processes are discussed in Sect. 3.2.5.

3.2.2 Nucleation and Growth of Surface Nuclei

An adion deposited on the surface of *a perfect crystal* stays on the surface as an adion only temporarily since its binding energy to the crystal is small. It is not a stable entity on the surface but it can increase its stability by the formation of clusters. The free energy of formation of a cluster of N ions, $\Delta G(N)$, has two components (terms)

$$\Delta G(N) = -Nze/\eta/ + \Phi(N) \qquad (30)$$

where the first term is related to the transfer of N ions from solution to the crystal phase and the second term is related to the increase of the surface energy owing to the creation of the surface of a cluster. This increase of the surface energy, or this excess energy, is equal to the difference of the binding energies of N bulk ions and N ions as arranged on the surface of the crystal.

Both the terms in Eq. (30) are functions of the size of the cluster N. The first term increases linearly with N and the second increases as $N^{2/3}$. The dependence of the energy of formation of a cluster $\Delta G(N)$ on the number of adions N in a 2D cluster is shown in Fig. 13. It can be seen that ΔG increases initially, reaches a maximum, and then decreases with increasing N. At the maximum the cluster size is N_c. The size of the critical nucleus N_c (the number of atoms in the cluster) in two-dimensional (2D) nucleation is given by

$$N_c = S\varepsilon/ze\eta \qquad (31)$$

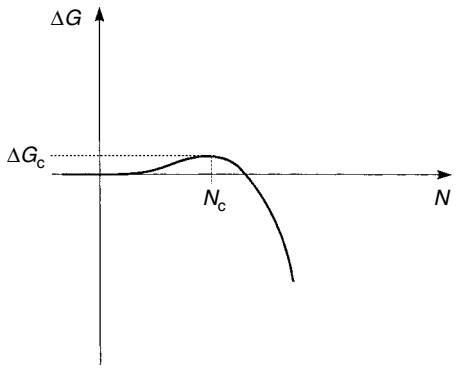


Fig. 13 Free energy of formation of a cluster as a function of size N (a cluster of N atoms); N_c, the size of the critical cluster (nucleus).

where b is the factor relating the surface area S of the nucleus to its perimeter P ($b = P^2/4S$; $b = \pi$ for a circular nucleus), s is the area occupied by an atom on the surface of the nucleus, and ε is the edge energy. For Ag nuclei, using Eq. (31) one calculates that the N_c at $\eta = 5$, 10, and 25 mV is 128, 32, and 5, respectively ($S = 6.55 \times 10^{-16}$ cm^2, $\varepsilon = 2 \times 10^{-13}$J cm^{-1}, $z = 1$, and $e = 1.602 \times 10^{-19}$ C). Thus, N_c *strongly depends on the overpotential*; it is inversely proportional to η^2. The critical radius of the surface nucleus r_c is a function of the overpotential

$$r_c = s\varepsilon/ze\eta \qquad (32)$$

For Ag ($r = 1.444$ Å), the critical radius r_c at $\eta = 5$, 10, and 25 mV is 16.35, 8.18, and 3.27 Å, respectively.

The spontaneous growth of clusters is possible after the maximum ΔG is reached. This critical cluster is the nucleus of the new phase and it is characterized by equal probability for growth and dissolution. The growth of clusters before a maximum is reached, and when average ΔG is increasing, is due to statistical energy fluctuations, which allow local higher values of ΔG, beyond the maximum [22].

Kolb et al. [18, 23, 24] formed nanoclusters of Cu on Au substrate by depositing Cu from solution onto the STM tip and subsequently transferring Cu from the tip to the substrate by an appropriate tip approach (Fig. 14). This technique can be fully automated and used for the fabrication of a variety of products in nano-electrochemical technology (e.g. conductors in microelectronics).

3.2.2.1 Nucleation of Surface Nuclei

The nucleation law for a uniform probability with time t of conversion of a site on the metal electrode into nuclei is given by

$$N = N_o[1 - \exp(-At)] \qquad (33)$$

where N_o is the total number of sites (the maximum possible number of nuclei on unit surface) and A is the nucleation rate constant. This equation represents the first-order kinetic model of nucleation. The rate of 2D nucleation J is given by

$$J = k_1 \exp[-bs\varepsilon^2/zekT\eta] \qquad (34)$$

where k_1 is the rate constant; b, the geometric factor depending on the shape of the 2D cluster, Eq. (32); s, the area occupied by an atom on the surface of the cluster; ε, the specific edge energy; k, the Boltzmann's constant; the other symbols have their usual meaning.

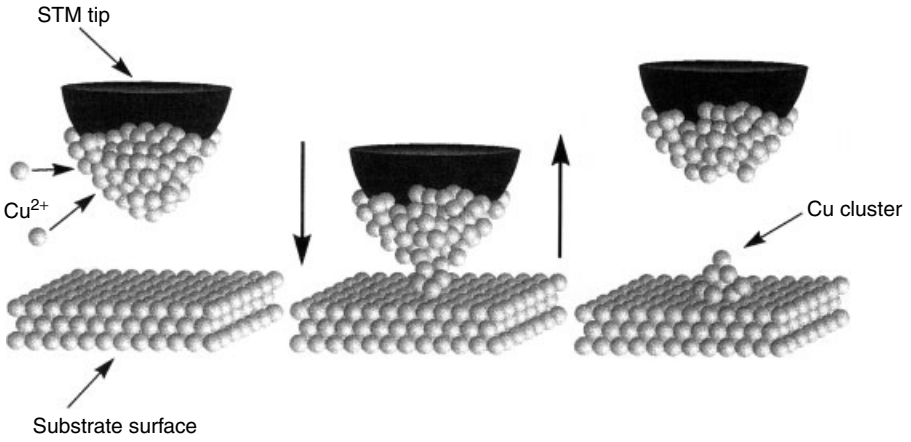


Fig. 14 Schematic diagram for the transfer of electrodeposited Cu from a Cu-covered STM tip to the Au substrate during a tip approach (from Ref. 18 with permission from Elsevier Science Inc.).

There are two limiting cases for Eq. (33) for the initial stages of nucleation (low t value). First, for a large nucleation constant A, Eq. (33) reduces to

$$N \approx N_o \tag{35}$$

indicating that all electrode sites are converted to nuclei instantaneously. Hence, this is referred to as *instantaneous nucleation*. Second, for small A and small t, Eq. (33) reduces to

$$N \approx AN_o t \tag{36}$$

since the exponential term in Eq. (33) can be represented as a linear approximation ($-e^{-At} \approx -1 - At$). In this case, the number of nuclei N is a function of time t and the nucleation is termed *progressive nucleation*.

It is possible to distinguish between these two modes of nucleation experimentally, for example, by the use of potentiostatic current–time transients (see Chapter 5 in Volume 2).

3.2.2.2 Growth of Surface Nuclei

In the initial stages of the growth of the nuclei, it can be assumed that nuclei grow independent of each other. This is valid only for the initial stages of growth. However, in the succeeding stages, it is necessary to consider the effect of the overlap between diffusion fields around growing nuclei [1]. The result of this overlap is the development of local

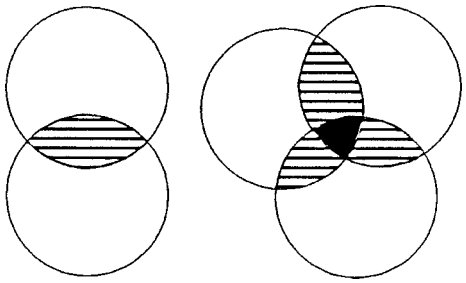


Fig. 15 Overlap of diffusion zones of cylindrical nuclei growing on a surface. Shaded regions, two zones overlapping; black region, three zones overlapping.

concentration and overpotential distribution in the neighborhood of the growing nuclei (clusters). Overlap areas are zones of reduced concentration and reduced nucleation rate (Fig. 15). Growing nuclei cannot grow freely in all directions since they will impinge on each other. The growth will stop at the point of contact, resulting in the limitation of the size of growth center. Theoretical analysis of potentiostatic current–time transients (Sect. 3.1.5) was used to formulate models of growth of surface nuclei [1, 25, 26]. For further discussion see Chapter 5 in Volume 2.

3.2.3 Layer Growth Mechanism

A schematic presentation of two basic mechanisms for the formation of a coherent deposit is shown in Fig. 16. In the layer growth mechanism, a crystal grows by the lateral spreading of discrete layers (steps) one after another across the surface. In this case, a growth layer, a step, is a structure component of a coherent deposit. Steps are structure components for the construction of a variety of growth forms in the electrodeposition of metals (e.g. columnar crystals, whiskers, fiber textures). We can distinguish between monoatomic steps, polyatomic microsteps, and polyatomic macrosteps.

3.2.3.1 Monoatomic Steps

Propagation of monoatomic steps (terraces) on the (111) Cu surface was monitored by AFM [20].

3.2.3.2 Microsteps

Propagation of copper microsteps on real surfaces of Pt was observed by Rynders and Alkire [27] using in situ AFM. They used the Pt single crystal as the substrate that was cut not exactly parallel to the (100) plane but with a misorientation of about 2°. This misorientation in cutting resulted in microsteps of 25 to 50 Å in height and terraces approximately 1 μ in length. Copper was electrodeposited from the $CuSO_4/H_2SO_4$ electrolyte. A sequence of potential (or current) pulses (10 ms to 10 s) was used to obtain a sequence of data that represents the initial growth processes. AFM images show that the

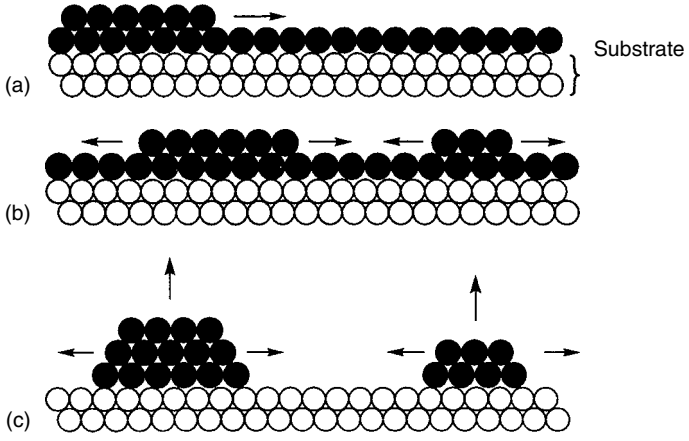


Fig. 16 Schematic representation of the layer growth (a,b) and the nucleation-coalescence mechanism (c).

growth of copper on Pt(100) at 1 mA cm^{-2} initiated along the largest of the steps and that the deposition rate was uniform along the step. At an intermediate current density, 10 mA cm^{-2}, deposit grew on all the steps. At current densities above 15 mA cm^{-2}, deposition occurred on both step and terrace sites.

Propagation of microsteps with a height of 30–100 Å (15–50 monoatomic layers) on a quasi-ideal surface of Ag was directly observed by Bostanov et al. using Nomarski differential contrast technique [28].

3.2.3.3 Macrosteps

Polyatomic macrosteps originate either from screw dislocations or from 3D nucleation. In the former case, steps are self-perpetuating and in the latter case, they are nucleation dependent. Macrosteps can be formed by the bunching of 1000 or more microsteps. There is, in general, the tendency of a large number of thin steps to bunch into a system of a few thick steps. Many monoatomic steps can unite (bunch, coalesce) to form a polyatomic step. Frank proposed the bunching mechanism to explain this process [29]. Bunching of steps is illustrated schematically in Fig. 17.

Wranglen observed using the optical microscope that a great many layers originate, one after the other, at growth centers [30]. At low current densities, a new layer does not start until the former has reached the edge of the crystal. At higher current densities, new layers are developed before the forerunners have reached their final size at the edge of the crystal. In this case, a great many layers propagate simultaneously. Damjanovic et al. observed in situ lateral and vertical growth of macrosteps using optical microscopy [31]. They found that the average velocity of the step motion is about 2×10^{-6} cm s^{-1} in the case of current density of 5 mA cm^{-2} for the electrodeposition of Cu on Cu(100) single-crystal face from $CuSO_4/H_2SO_4$ solution. The average height of steps varies linearly with the thickness of the deposit ($2–12 \times 10^{-6}$ cm at 5 mA cm^{-2} for the time of deposition of 9–35 min).

Giesen et al. reported an analysis of step fluctuations on a Cu(100) electrode in aqueous electrolyte using STM [19]. These fluctuations involve a rapid exchange of atoms along the steps, between neighboring steps, and on adjacent terraces. They determined the mass transport involved in the fluctuations as a function of the electrode potential.

3.2.3.4 Lateral Merging of Steps

Bertocci and Bertocci found that the principal mechanism for the formation of the defect clusters in the electrodeposition of Cu on single crystals is the interaction between steps laterally merging and annihilating [32]. Defect concentration tends to

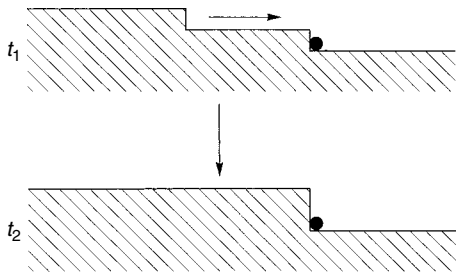


Fig. 17 Bunching of steps.

increase when the misorientation is larger than a few degrees.

3.2.4 Nucleation-coalescence Growth Mechanism

In this case, the structure components are 3D crystallites, and a coherent deposit is built by the coalescence (joining) of these crystallites. As an example of nucleation-coalescence growth mechanism, we describe the nucleation and growth of electrodeposited gold on (111) surfaces of silver by means of electron microscopy. Dickinson et al. [33] found that growth from $HAuCl_4-KCl$ solutions occurs by a nucleation mechanism and that in the first stage of deposition gold is in the form of very thin platelike isolated islands (or three-dimensional crystallites, TDCs). At this stage, the isolated nuclei or TDC had a population density of $j4 \times 10^{10}$ cm^{-2} and the total surface coverage was about 22%. After this stage of isolated nuclei, further deposition leads to the coalescence of some of the TDCs into the form of elongated crystallites with a reduction in population to $j1 \times 10^{10}$ cm^{-2}. In the next stage, a linked network structure is formed, corresponding to a surface coverage of about 58%. The next structure can be described as a continuous deposit containing holes and channels. Further deposition leads to the filling of these holes and channels and a complete hole-free film is formed at a thickness of 80–100 Å.

Thus, the sequence of growth of electrodeposited gold by nucleation-coalescence mechanism has four stages: (1) formation of isolated nuclei and their growth to TDC, (2) coalescence of TDC, (3) formation of linked network, and (4) formation of a continuous deposit.

It is interesting to note that it was concluded in this work that the first few angstroms of deposit form an alloy layer. Another interesting and important conclusion in this work is that the sequence of growth of electrodeposits (at 20 °C) is very similar to that of the evaporated deposits (at 250–300 °C).

Three-dimensional epitaxial crystallites (TECs) were observed in the first stages of electrodeposition of copper [34] and nickel [35] on copper substrates. TEC of nickel formed on copper-film substrate from the nickel sulfate solutions in low concentrations are rectangular in shape with the edges having an average length of 1300 Å. The coherent deposit was formed by the growth and coalescence of TEC. The continuous deposit was observed when the average thickness was 20–50 Å.

3.2.4.1 Coalescence-induced Defects

The two most significant coalescence-induced defects are voids and dislocations. Nakahara [36] has shown by transmission electron microscopy that microscopic voids (<50 Å) are generated in electrodeposited Cu, Au, and Ni–P films at the boundaries between 3D crystallites during their coalescence. When the coalescing crystallites are misoriented with respect to each other, the coalescing boundaries are grain boundaries. Weil and Wu [37] have shown that misorientation (misalignment) between neighboring Ni crystallites generate dislocations.

3.2.4.2 Deposition of a Metal on a Foreign Metallic Substrate

In the industrial applications of metal deposition, a metal M is deposited either on the native metal substrate M or on a foreign metal substrate S. As an example of the former, Cu is electrodeposited on a Cu substrate formed by electroless Cu

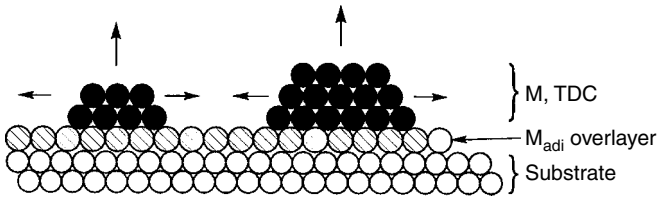


Fig. 18 Stranski-Krastanov mechanism.

deposition on an activated nonconductor in the fabrication of printed circuit boards. As an example of the latter, Ni is electrodeposited on Cu in the fabrication of contact pads in the electronics industry.

The mechanism of growth of metal M on a foreign metallic substrate S is determined by the two most important parameters: (1) the binding energy between M and M and M and S and (2) the crystallographic misfit between M and S.

Three mechanisms are possible in this case and they are related to the two above-mentioned parameters. The first mechanism is the *TDC nucleation and growth* or *Volmer–Weber mechanism*. This mechanism is operative when the binding energy between adions of metal M (M_{adi}) and the substrate M, $\psi(M_{adi} - M)$, is larger than that between M_{adi} and metal S, $\psi(M_{adi} - S)$

$$\psi(M_{adi} - M) \gg \psi(M_{adi} - S) \qquad (37)$$

In this case, the crystallographic misfit does not have any effect. The crystallographic misfit *mf* is defined by

$$mf = \frac{a_S - a_M}{a_M} \qquad (38)$$

where a_S and a_M are the lattice spacing for the substrate S and deposit M respectively.

One common characteristic of mechanisms 2 and 3 is the relationship between the binding energies, and the difference is in the misfit.

Thus, if

$$\Psi(M_{adi} - S) \gg \psi(M_{adi} - M) \qquad (39)$$

we distinguish two types of mechanisms:

1. If $a_S \approx a_M$, misfit is zero, the deposition mechanism is *layer-by-layer* or the *Frank-Van der Merve growth mechanism*.
2. If $a_S \neq a_M$, misfit is present, positive or negative, then the growth proceeds by *the Stranski-Krastanov mechanism*, which is composed of two steps. In the first step, a 2D overlayer of M_{adi} on S is formed and in the second step TDCs grow on top of this predeposited overlayer (Fig. 18).

3.2.5 Effect of Additives

Most solutions used in the electrodeposition of metals and alloys contain one or more inorganic or organic additives that have a specific function in the deposition process. These additives (and impurities) affect the deposition and crystal-building processes as adsorbates (substances that are adsorbed) at the surface of the cathode. Adsorption and the factors that determine the adsorbate–surface interaction are discussed in detail in Volume 1 and Volume 2.

3.2.5.1 Chemisorption and Physisorption

One classification of adsorption phenomena is based on the adsorption energy, that

is, the energy of the adsorbate–surface interaction. In this classification there are two basic types of adsorption: chemisorption (an abbreviation of chemical adsorption) and physisorption (an abbreviation of physical adsorption). In chemisorption, the chemical attractive forces of adsorption act between the surface and the adsorbate (usually covalent bonds). Thus, there is a chemical combination between the substrate and the adsorbate where electrons are shared and/or transferred. New electronic configurations are formed by this sharing of electrons. In physisorption, the physical forces of adsorption, van der Waals or electrostatic forces, operate between the surface and the adsorbate; there is no electron transfer and no electron sharing.

Adsorption energy for chemisorbed species is larger than that for physisorbed species. Typical values for chemisorption are in the range of 20–100 kcal mol^{-1} and for physisorption in the range of 5 kcal mol^{-1}.

3.2.5.2 Adsorption Equilibrium

Since in many cases, electrodeposition in the presence of an additive does not use up the additive (no incorporation of the additive in the deposit), one can conclude that the adsorption equilibrium is a dynamic one. In a dynamic adsorption equilibrium state, the adsorbed molecules continually desorb at a rate equal to the rate at which dissolved molecules from the solution become adsorbed. If the rates of the adsorption and desorption processes are high and of the same order of magnitude as that of the cathodic deposition process, then no incorporation, entrapment, of additives in the deposit will occur. However, if they are much smaller, additive molecules will be entrapped in the deposit via propagating steps (growing crystallites). Thus, at a current density higher than the optimum value, additives (brighteners or levelers) will be incorporated into the deposit. This incorporation can result in poor quality of the resulting deposit.

3.2.5.3 Effect of Additives on Kinetics and Mechanism of Electrodeposition

In the discussion of atomistic aspects of electrodeposition of metals, Sect. 3.1.5, it was shown that in the electrodeposition the transfer of a metal ion M^{n+} from the solution into the ionic metal lattice in the electrodeposition process may proceed via one of the following two mechanisms: (1) direct mechanism in which ion transfer takes place on a kink site of step edge or on any site on the step edge (any growth site); and (2) terrace site ion mechanism. The adsorbed additives affect both these mechanisms by changing the concentration of growth sites c_{gs} on the surface (n_{gs}/cm^2, n_{gs} – number of growth sites), concentration of adions, c_{adi} on the surface, diffusion coefficient D_{adi}, and the activation energy E_{adi} of surface diffusion of adions.

3.2.5.4 Effect of Additives on Nucleation and Growth

In the presence of adsorbed additives, the mean free path for lateral diffusion of adions is shortened, which is equivalent to a decrease in the diffusion coefficient D (diffusivity) of adions. This decrease in D can result in an increase in adion concentration at steady state and, thus, an increase in the frequency of the two-dimensional nucleation between diffusing adions. Additives can also influence the propagation of microsteps and cause bunching and formation of macrosteps (Fig. 17).

3.2.5.5 Dependence of Types of Deposit on the Surface Coverage by Additive

The type of deposit obtained at constant current density may depend on the value of the surface coverage θ by an additive. Damjanovic et al. [31] studied the effect of various values of the surface coverage θ of *n*-decylamine on the growth form of copper on the (100) plane of copper single crystal at 5 mA cm^{-2}. The surface coverage θ was varied by the addition of a known amount of *n*-decylamine to a highly purified solution of $CuSO_4$, H_2SO_4, and H_2O. The coverage θ of *n*-decylamine was estimated from the adsorption isotherm for *n*-decylamine on copper in 1 N $NaClO_4$. It was found that when $\theta < 10^{-2}$ (at the bulk concentration of *n*-decylamine $<10^{-3}$ mol L^{-1}) the crystal growth form was of the layers type. At $\theta \geq 10^{-2}$, the deposit observed was of the ridge type. Thus, variation in the surface coverage θ resulted in two entirely different types of deposit.

3.2.5.6 Leveling

Leveling was initially defined as the progressive reduction of the surface roughness during deposition. Surface roughness may be the result of a coarse mechanical polishing. In this case, scratches on the cathode represent the initial roughness, and the result of cathodic leveling is a smooth (flat) deposit or a deposit of reduced roughness. During this type of leveling, more metal is deposited in recesses than on peaks. This leveling is of great value in the metal-finishing industry. It is also of great importance in the electronics industry for the deposition through polymeric masks and in the electroforming of micromechanical devices (sensors, actuators, micromotors). In these cases, the leveling process means the uniform deposition in microprofiles (filling up recesses), which can be defined with masks or without. Leveling can be achieved in solutions in the absence of addition agents and in the presence of the same leveling agents. Thus, there are two types of leveling processes: (1) geometric leveling corresponding to leveling in the absence of specific agents and (2) true or electrochemical leveling corresponding to leveling in the presence of leveling agents [1].

Dukovic and Tobias [38] and Madore and Landolt [39] developed theoretical models of leveling during electrodeposition in the absence and in the presence of additives. Both the models show that a significant leveling of semicircular and triangular grooves by uniform current distribution (geometric leveling) is achieved when the thickness of the deposit is at least equal to the depth of the groove. We show in the next section that leveling in the presence of leveling agents (true leveling) is achieved much earlier.

Theories of leveling by additives are based on (1) the correlation between an increase in polarization produced by the leveling agents and (2) the preferential adsorption of a leveling agent on the high point (peaks or flat surfaces), Kardos and Foulke [40].

3.2.5.7 Brightening

Kardos and Foulke [40] distinguish three possible mechanisms for bright deposition: (1) diffusion-controlled leveling, (2) grain refining, and (3) randomization of crystal growth on favorable sites. Some fundamental aspects of brightening and leveling are reviewed by Onicin and Muresan [41].

3.2.5.8 Consumption of Additives

Additives can be consumed at the cathode by incorporation into the deposit and/or by the electrochemical reaction at the cathode or anode. Consumption of coumarin

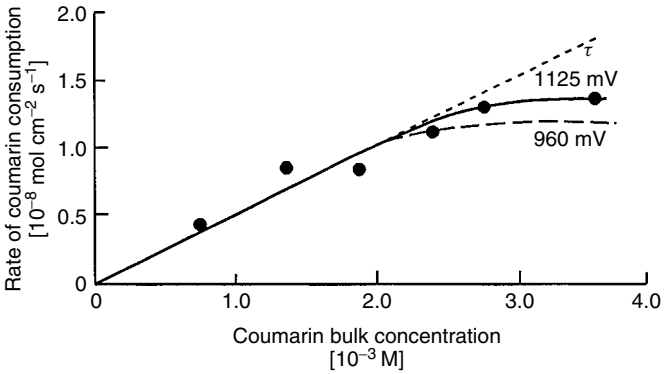


Fig. 19 Rate of coumarin consumption as a function of coumarin concentration; constant potential experiments: −960 and −1125 mV (vs saturated calomel electrode, SCE), 985 rpm, pH 4; τ, theoretical rate transport (from Ref. 42 (1968) with permission from Elsevier Science Inc.).

in the deposition of nickel from a Watts-type solution was studied extensively; for example, Roger and Taylor found that the coumarin concentration decreases linearly with time and that the rate of coumarin consumption is a function of coumarin bulk concentration [42]. Figure 19 shows that the rate of consumption increases with an increase in the bulk concentration of coumarin. Thus, in the case of the electrodeposition of nickel from Watts-type solution the total current density is the sum of the current densities for nickel deposition i_{Ni}, hydrogen evolution i_H, and additive reduction i_R.

In general, control of an electrodeposition solution includes monitoring the concentration of additives and by-products of the reaction of additives at the electrode.

3.3 Electroless Deposition

3.3.1 Introduction

The basic components of an electrolytic cell for the electrodeposition of metals from an aqueous solution are power supply, two metal electrodes (M_1 and M_2), water containing the dissolved ions, and two metal–solution interfaces; M_1–solution and M_2–solution. An electrolytic cell for the electroless deposition is shown in Fig. 20. It can be seen that in electroless deposition there is no power supply, and the system has only one electrode. However, the solution is more complex. It contains water, a metal salt MA (M^{z+}; A^{z-}), and a reducing agent Red as basic components.

The overall reactions of electrodeposition and electroless deposition may be used to compare these two processes. The process of *electrodeposition* of metal M is represented by

$$M^{z+}_{solution} + ze \xrightarrow{\text{Electrode}} M_{lattice} \qquad (40)$$

In this process, z electrons are supplied by an external power supply. The overall reaction of *electroless* metal deposition is

$$M^{z+}_{solution} + Red_{solution} \xrightarrow{\text{Catalytic surface}} M_{lattice} + Ox_{solution} \qquad (41)$$

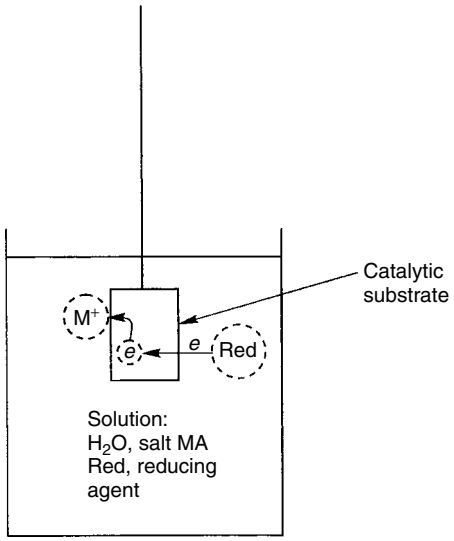


Fig. 20 An electrolytic cell for electroless deposition of metal M from an aqueous solution of metal salt MA and a reducing agent Red.

where Ox is the oxidation product of the reducing agent Red. The catalytic surface may be the substrate S itself or the catalytic nuclei of metal M dispersed on a noncatalytic substrate surface. In the electroless deposition process, a reducing agent Red in the solution is the electron source; the electron-donating species Red gives electrons to the catalytic surface and metal ions M^{z+} at the surface. The reaction represented by Eq. (41) must be conducted in such a way that a homogeneous reaction between M^{z+} and Red, in the bulk of the solution, is suppressed. Another comparison between electrodeposition and electroless deposition points to more basic differences. In electrodeposition, there are two electrodes: a cathode and an anode. Here, two separate electron-transfer reactions occur at two spatially separated electrode–electrolyte interfaces. At the cathode, the reduction reaction occurs, Eq. (40), and at the anode, an oxidation reaction occurs, for example,

$$M_{lattice} \longrightarrow M^{2+} + ze \qquad (42)$$

In the electroless deposition, the two electrochemical reactions, the reduction of $M^{z+}_{solution}$ and the oxidation of $Red_{solution}$ occur at the same electrode, at one and the same electrode–electrolyte interface (Eq. (41); Fig. 20). Thus, in electroless deposition, there is a statistical division of the catalytic sites on the substrate into anodic and cathodic sites. Since these catalytic sites are part of the same piece of metal (substrate), there is a flow of electrons between these sites.

In this section, we discuss the electrochemical model of electroless deposition, kinetics and mechanism of partial reactions, activation of noncatalytic surfaces, kinetics of electroless deposition, and the mechanism of electroless crystallization.

3.3.2 Electrochemical Model: the Mixed-potential Theory

An electrochemical model for the process of electroless metal deposition was suggested by Paunovic [43] and Saito [44] based on the Wagner and Traud [45] mixed-potential theory of corrosion processes. According to the mixed-potential theory of electroless deposition, the overall reaction given by Eq. (41) can be decomposed into one reduction reaction, the cathodic partial reaction

$$M^{z+}_{solution} + ze \xrightarrow{\text{Catalytic surface}} M_{lattice} \qquad (43)$$

and one oxidation reaction, the anodic partial reaction

$$Red_{solution} \xrightarrow{\text{Catalytic surface}} Ox_{solution} + me \qquad (44)$$

Thus, the overall reaction (Eq. (41)) is the outcome of the combination of two different partial reactions, Eqs (43 and 44). As mentioned above, these two partial reactions, however, occur at one electrode, the same metal–solution interphase. The equilibrium (rest) potential of the reducing agent, $E_{eq,Red}$ (Eq. 34) must be more negative than that of the metal electrode, $E_{eq,M}$ (Eq. 33), in order that the reducing agent Red can function as an electron donor and M^{z+}, as an electron acceptor. This is in accord with the general fact that an electrode with lower electrode potential will reduce ions of an electrode with higher electrode potential. A high positive standard electrode potential indicates a strong tendency toward reduction; a low negative standard electrode potential indicates a strong tendency toward the oxidized state.

3.3.2.1 Wagner–Traud Diagram

According to the mixed-potential theory, the overall reaction of the electroless deposition, Eq. (31), can be described electrochemically in terms of three current–potential ($i-E$) curves, as shown schematically in Fig. 21. First, there are two current–potential curves for the partial reactions (solid curves): (1) $i_c = f(E)$, the current–potential curve for the reduction of M^{z+} ions, recorded from the rest potential $E_{eq,M}$, in the absence of the reducing agent Red (when the activity of M^{z+} is equal to 1 then $E_{eq,M} = E^o M$) and (2) $i_a = f(E)$, the current–potential curve for the oxidation of the reducing agent Red, recorded from the rest potential $E_{eq,Red}$, in the absence of M^{z+} ions (when the activity of Red is equal to 1 then $E_{eq,Red} = E^o Red$). Then, the third curve, $i_{total} = f(E)$, the dashed curve in Fig. 21, is the current–potential curve for the overall reaction.

The two major characteristics of this system of curves are as follows:

1. $i_{total} = f(E)$ curve intersects the potential axis. At this intersection, the current

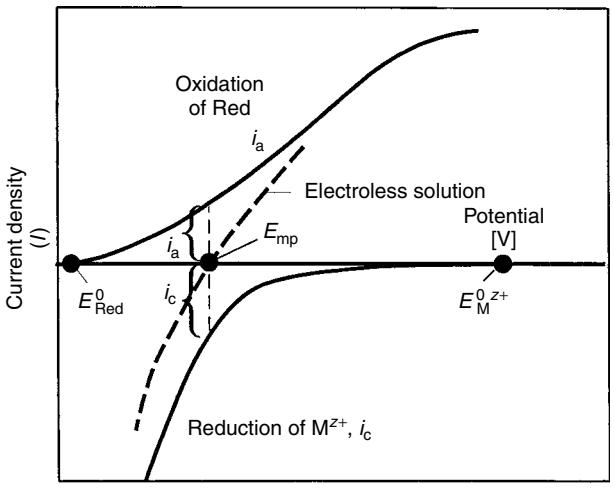


Fig. 21 Wagner–Traud diagram for the total (i_{total}) and component current–potential curves (i_a, i_c) for the overall reaction of electroless deposition.

is zero and

$$i_c = i_a \text{ for } i_{total} = 0 \qquad (45)$$

The potential where Eq. (45) holds is the mixed potential, E_{mp}.

2. At any point of the $i_{total} = f(E)$ curve

$$i_{total} = i_a + i_c \qquad (46)$$

Thus, the total current density, i_{total}, is composed of two components. It is a result of the addition of current densities of the two partial processes.

3.3.2.2 Evans Diagram

An alternative method of presenting the current–potential curves for electroless metal deposition is the Evans diagram. In this method, the sign of the current density is suppressed. Figure 22 shows a general Evans diagram with current–potential functions $i = f(E)$ for the individual electrode processes, Eqs (43 and 44). According to this presentation of the mixed-potential theory, the current–potential curves for individual processes, $i_c = i_M = f(E)$ and $i_a = i_{Red} = f(E)$, intersect. The coordinates of this intersection have the following meaning: (1) the abscissa, the current density of the intersection, is the deposition current density i_{dep} (i.e. $\log i_{dep}$), which is the rate of electroless deposition in terms of milliamperes per square centimeter and (2) the ordinate, the potential of the intersection is the mixed potential, E_{mp}.

3.3.2.3 Mixed Potential, E_{mp}

When a catalytic surface S is introduced into an aqueous solution containing M^{z+} ions and a reducing agent, the partial reaction of reduction, Eq. (43), and the partial reaction of oxidation, Eq. (44), occur simultaneously. Each of these partial reactions strives to establish its own equilibrium, E_{eq}. The result of these processes is a steady state with the compromised potential called the *steady-state mixed potential*, E_{mp}. The result of this mixed potential is that the potential of the redox couple Red/Ox, Eq. (44), is raised anodically from the reversible value $E_{eq,Red}$ (Fig. 22) and the potential of the metal electrode M/M^{z+}, Eq. (43), is depressed cathodically

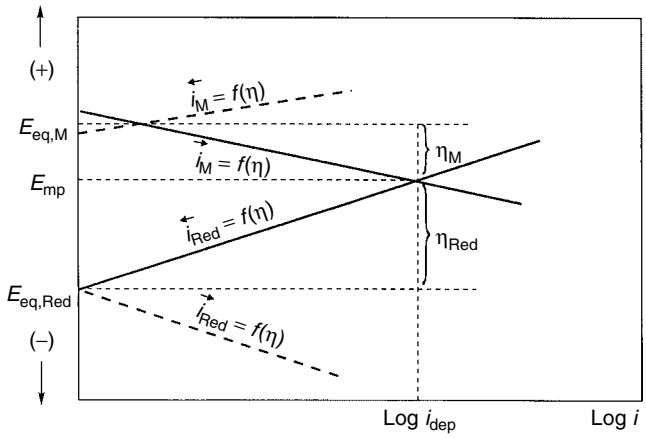


Fig. 22 Evans diagram of current–potential curves for the system with two different simultaneous electrochemical reactions. Kinetic scheme: Eqs (43 and 44).

from its reversible value $E_{eq,M}$, down to the mixed potential E_{mp} (Fig. 22).

Thus, the four basic characteristics of the steady-state mixed potential are as follows:

1. Both redox systems are shifted from their own characteristic equilibrium potentials by the amount η (overpotential)

$$\eta_M = E_{mp} - E_{eq,M} \quad (47)$$

$$\eta_{Red} = E_{mp} - E_{eq,Red} \quad (48)$$

2. A net electrochemical reaction occurs in each redox system since both reactions, Eqs (43 and 44), are shifted from their equilibrium by the introduction of the mixed potential.
3. The condition for steady state is that the rate of reduction of M^{z+}, the cathodic current density i_M, is equal to the rate of oxidation of the reducing agent Red, the anodic current density i_{Red}

$$i_{M,deposition} = (i_M)E_{mp} = (i_{Red})E_{mp} \quad (49)$$

since a net current cannot flow in the isolated system.

4. A system at the steady-state mixed potential is not in equilibrium since a net overall reaction does occur, and therefore, change of the free energy is not zero, which is required for thermodynamic equilibrium.

Figure 23 shows an experimental example of Evans diagram for electroless deposition of copper.

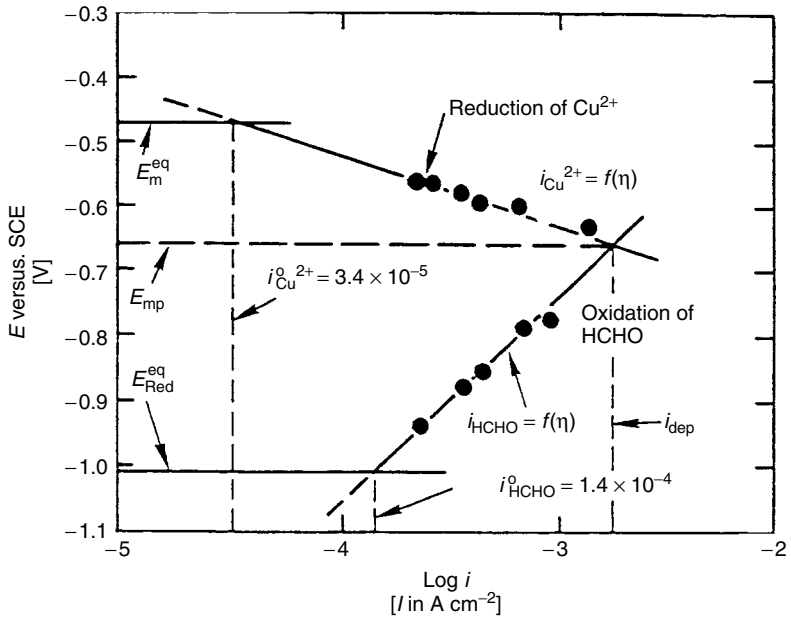


Fig. 23 Evans diagram for electroless deposition of copper. Current–potential curves for the reduction of Cu^{2+} ions and for the oxidation of reducing agent Red, formaldehyde, combined into one graph. Solution for the Tafel line for the reduction of Cu^{2+} ions: 0.1 M $CuSO_4$, 0.175 M EDTA, pH 12.50, E_{eq} (Cu/Cu^{2+}) = −0.47 V versus SCE; for the oxidation of formaldehyde: 0.05 M HCHO and 0.075 M EDTA, pH 12.50, E_{eq}(HCHO) = −1.0 V versus SCE; temperature 25 °C (±0.5 °C) (from Ref. 43 with permission from American Electroplaters and Surface Finishers Society).

The validity of the mixed-potential theory is demonstrated for the electroless deposition of Cu [44, 46–48], Ni [49], and Au [50].

3.3.3 Kinetics and Mechanism

3.3.3.1 The Cathodic Partial Reaction. Kinetic Scheme

Metal ions in a solution for electroless metal deposition have to be, in general, complexed with a ligand. Complexing is necessary in order to prevent formation of metal hydroxide, for example $Cu(OH)_2$, in the electroless copper deposition. One of the fundamental problems in the electrochemical deposition of metals from complexed ions is the problem of electroactive (charged) species. The electroactive species may be complexed or noncomplexed metal ion [1]. In the first case, the kinetic scheme of the process of metal deposition is one of simple charge transfer. In the second case, the kinetic scheme is that of charge transfer preceded by the dissociation of the complex. The mechanism of the second case involves a sequence of at least two basic elementary steps:

1. formation of the electroactive species and
2. charge transfer from the catalytic surface to the electroactive species.

Electroactive species M^{z+} are formed in the first step by the dissociation of the complex $[ML_x]^{z+xp}$

$$[ML_x]^{z+xp} \longrightarrow M^{z+} + xL^{p} \qquad (50)$$

where p is the charge state of the ligand L, z is the charge of the noncomplexed metal ion, and $(z + xp)$ is the charge of the complexed metal ion.

The charge transfer

$$M^{z+} + ze \longrightarrow M \qquad (51)$$

proceeds in steps, usually with the first charge transfer (one-electron transfer) serving as the RDS

$$M^{z+} + e \xrightarrow{\text{RDS}} M^{(z-1)+} \qquad (52)$$

Thus, from the kinetic aspects, the cathodic partial reaction is an electrochemical reaction, Eq. (51), which is preceded by a chemical reaction, Eq. (50).

3.3.3.2 Kinetics

The major factors determining the rate of the partial cathodic reaction are concentrations of metal ions and ligands, pH of the solution, and the type and concentration of additives. These factors determine the kinetics of the partial cathodic reaction in a general way, as given by the fundamental electrochemical kinetic equations discussed in Sect. 3.1. Additives may have two opposing effects: acceleration and inhibition [51]. The major factors that are most likely responsible for the acceleration effect of additives are (1) the charge density of the electron system of the additive and (2) the exchange of electrons between electrode, O-bonded additive molecule, and the complexed metal ions in the solution. Inhibition effect and the ''cathodic passivation'' are explained in terms of blocking of the catalytic surface which results in the decrease in the available surface area [52].

3.3.3.3 The Anodic Partial Reaction. Mechanism

The overall anodic partial reaction, Eq. (44), usually proceeds in at least two elementary steps (like the cathodic partial reaction): the formation of the electroactive species and charge transfer.

The formation of electroactive species (R) proceeds usually in two steps, through an intermediate (Red_{interm})

$$Red \longrightarrow Red_{interm} \quad (53)$$

$$Red_{interm} \longrightarrow R \quad (54)$$

Van den Meerakker [53] proposed the following general mechanism for the formation of electroactive species R from the intermediate Red_{interm}, represented now by R−H

$$Red_{interm} = R{-}H \longrightarrow R_{ads} + H_{ads} \quad (55)$$

where R_{ads} is the electroactive species R adsorbed at the catalytic surface. According to this mechanism, Eq. (55), the electroactive species R_{ads} is formed in the process of dissociative adsorption (dehydrogenation) of the intermediate R−H that involves breaking R−H bond.

The adsorbed hydrogen, H_{ads}, may be desorbed in the chemical reaction

$$H_{ads} \longrightarrow \tfrac{1}{2}H_2 \quad (56)$$

or in the electrochemical reaction

$$H_{ads} \longrightarrow H^+ + e \quad (57)$$

In the electroless deposition of copper, for example, when the reducing agent is formaldehyde and the substrate is Cu, H_{ads} desorbs in the chemical reaction, Eq. (56). If the substrate is Pd or Pt, hydrogen desorbs in the electrochemical reaction, Eq. (57).

The most studied anodic partial reaction is the oxidation of formaldehyde, Red = H_2CO. The overall reaction of the electrochemical oxidation of formaldehyde at the copper electrode in an alkaline solution proceeds according to the equation

$$H_2CO + 2OH^- \longrightarrow HCOO^- + H_2O + \tfrac{1}{2}H_2 + e \quad (58)$$

The mechanism of this reaction involves the following sequence of elementary steps [53, 54]: (1) formation of electroactive species $[HC(OH)O^-]_{ads}$ in three steps and (2) charge transfer: the electrochemical oxidation (desorption) of electroactive species R_{ads}

$$[HC(OH)O^-]_{ads} + OH^- \longrightarrow HCOO^- + H_2O + e \quad (59)$$

where $HCOO^-$ is the oxidation product of R_{ads}(Ox).

A similar kinetic scheme can be applied to other reducing agents, for example, borohydride (Red = BH_4^-), hypophosphite ($H_2PO_2^-$), hydrazine (NH_2NH_2) where the electroactive species RH are $[BH_2OH^-]_{ads}$, $[HPO_2^-]_{ads}$, and $[N_2H_3]_{ads}$, respectively [50, 53].

3.3.3.4 Parallel Reactions

There are parallel reactions in some cases of oxidation of the reducing agents; for example, in the case of oxidation of BH_4^- and $H_2PO_2^-$; parallel reactions are (probably cathodic reactions) resulting in the incorporation of B and P into the metal deposit, respectively. Thus, when electroless Ni is deposited from solutions containing BH_4^- as the reducing agent, we designate this deposit as Ni(B), and when the reducing agent is $H_2PO_2^-$, we designate the deposit as Ni(P).

3.3.3.5 Kinetics

The major factors determining the rate of the anodic partial reaction are pH and additives. Since OH^- ions are reactants in the charge-transfer step, for example, Eq. (59), the effect of pH is direct and significant (e.g. see Ref. 55). Additives may have an inhibiting or an accelerating effect.

3.3.4 Activation of Noncatalytic Surfaces

Noncatalytic surfaces (e.g. nonconductors, noncatalytic metals, noncatalytic semiconductors) have to be activated, that is, made catalytic, prior to the electroless deposition. This activation is performed by generating catalytic nuclei on the surface of a noncatalytic material. Two major types of processes have been used to produce catalytic nuclei: electrochemical and photochemical.

3.3.4.1 Electrochemical Activation

In the electrochemical method, catalytic nuclei of metal M on a noncatalytic surface S may be generated in an electrochemical oxidation–reduction reaction

$$M^{z+} + \text{Red} \longrightarrow M + \text{Ox} \qquad (60)$$

where M^{z+} is the metallic ion and M is the metal catalyst. The preferred catalyst is Pd and, thus, the preferred nucleating agent M^{z+} is Pd^{2+} (from $PdCl_2$). The preferred reducing agent Red in this case is Sn^{2+} ion (from $SnCl_2$). In this example, the overall reaction of activation, according to a simplified model, is

$$Pd^{2+} + Sn^{2+} \longrightarrow Pd + Sn^{4+} \qquad (61)$$

Sn^{2+} can reduce Pd^{2+} ions, since the standard oxidation–reduction potential of Sn^{4+}/Sn^{2+} is 0.15 V and that of Pd^{2+}/Pd is 0.987 V. The flow of electrons is from the more electronegative (here Sn^{4+}/Sn^{2+}) toward the less electronegative (more positive) couple (here Pd^{2+}/Pd). Since the standard potentials of Au^{+}/Au is 1.692 V, Sn^{2+} ions can reduce Au^{+} to produce Au catalytic nuclei.

Electrochemical activation using $PdCl_2$ and $SnCl_2$ may be performed either in two steps or in one step [1].

Some nonconductors, for example polymers like polycarbonates and polystyrenes, must be subjected to a surface treatment prior to activation in order to get good adhesion of palladium nuclei. Surface treatment can include the use of chemical etchings for plastics or reactive gas plasma treatments [56].

3.3.4.2 Photochemical Activation

Photochemical processes have been used to produce catalytic metallic nuclei in three ways, each characterized by a different kinetic scheme: photoelectrochemical, photoelectron, and intramolecular. These processes have been reviewed by Paunovic [57] and Zhang et al. [58] with numerous references.

3.3.4.3 Activation by Displacement Deposition

Silicon can be made catalytic for the electroless deposition of Ni by replacing the surface Si atoms with Ni atoms [59]

$$2Ni^{2+} + Si \longrightarrow 2Ni + Si^{4+} \qquad (62)$$

This reaction is called *displacement deposition* since the nickel ions in solution simply displace the silicon at the surface. The substrate, Si, acts here as a reducing agent as will be discussed in Sect. 3.4. Copper may be deposited on Si from HF acid solutions [60]. In the presence of HF, Si is oxidized into $[SiF_6]^{2-}$.

Similarly, aluminum substrates can be activated by a displacement reaction [1]

$$3Pd^{2+} + 2Al \longrightarrow 3Pd + 2Al^{3+} \qquad (63)$$

3.3.4.4 Activation by Thermal Decomposition of Metallic Oxides

The surface of alumina, Al_2O_3, may be activated by employing laser or UV irradiation to decompose Al_2O_3 [61]. Decomposition

of Al_2O_3 results in the generation of aluminum particles, which are catalytic for the electroless deposition of Cu (the first reaction probably being displacement deposition).

3.4 Displacement Deposition

3.4.1 Introduction

In Sect. 3.3, we have shown that in the case of electroless deposition the reducing agent Red in the solution is the electron source, the electron-donating species that give electrons to the catalytic surface and the metal ions M^{z+} at the interface. In this section, we show that the substrate itself can also be the electron-donating species.

Also, we show that, in general, Ox/Red (M^{z+}/M) couples with high standard electrode potentials are reduced by Ox/Red (M^{z+}/M) couples with low standard electrode potentials. In other words, low reduces high.

Consider a strip of Zn placed in a solution of $CuSO_4$ (Fig. 24). Consider the reactions that will occur in this system if the activity of Cu^{2+} ions is 1. This problem can be solved by considering standard electrode potentials. The standard electrode potential for Cu/Cu^{2+} is $E^o = 0.337$, and that for Zn/Zn^{2+} is $E^o = -0.763$. Since Zn/Zn^{2+} has a lower electrode potential than the Cu/Cu^{2+} system, Zn will reduce Cu^{2+} ions in the solution. Thus, the partial reactions in the system shown in Fig. 25 are

$$Zn \longrightarrow Zn^{2+} + 2e \qquad (64)$$

and

$$Cu^{2+} + 2e \longrightarrow Cu \qquad (65)$$

and the overall reaction is

$$Zn + Cu^{2+} \longrightarrow Zn^{2+} + Cu \qquad (66)$$

Thus, a layer of metallic Cu is deposited on the zinc, while Zn dissolves into solution. Metallic Zn under these conditions reduces the Cu^{2+} ions. This reaction is called a *displacement deposition of Cu on Zn*.

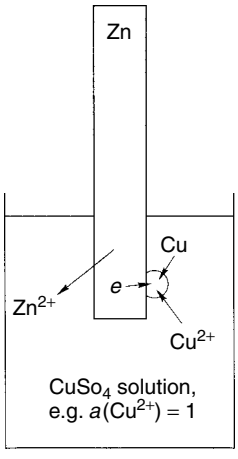


Fig. 24 A displacement deposition of Cu on Zn.

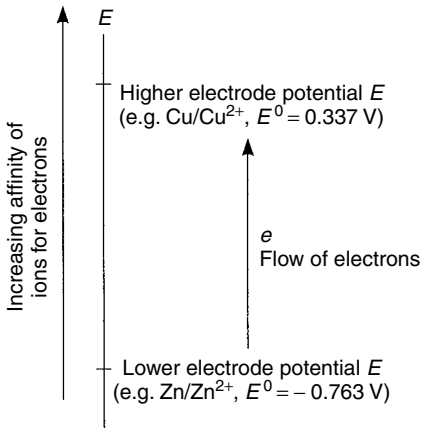


Fig. 25 An electrode with lower electrode potential will reduce ions of an electrode with higher electrode potential.

The thickness of the deposited metal in the displacement deposition is self-limiting, since the displacement deposition process needs exposed (free) substrate surface in order to proceed.

This technique comes in a variety of names, depending on the application: for example, immersion deposition, galvanic deposition (galvanic corrosion), conversion, and cementation (in the metal recovery industry).

3.4.2 Electrochemical Model

The overall displacement deposition reaction in general Ox/Red (M^{z+}/M) terms is given by

$$M_l + M_r^{z+} \longrightarrow M_r + M_l^{z+} \qquad (67)$$

for the cell

$$M_l/M_l^{z+}(aq)//M_r^{z+}(aq)/M_r \qquad (68)$$

The partial (half-cell) reactions are

$$\text{oxidation} : M_l \longrightarrow M_l^{z+} + ze \qquad (69)$$

$$\text{reduction} : M_r^{z+} + ze \longrightarrow M_r \qquad (70)$$

Metal substrate M_l dissolves into the solution, Eq. (69), and, thus, supplies the electrons necessary for the reduction, deposition reaction, Eq. (70). The relationship between the partial reactions presented above is shown in Fig. 26.

We have described one example of this type of electrochemical deposition in Sect. 3.4.1, when we considered processes at a strip of Zn placed in a solution of $CuSO_4$ (Fig. 25). We have stated that there are two partial reactions in that system, like in an electroless system. In the displacement deposition of Cu on Zn, electrons are supplied in the oxidation reaction of Zn, reaction (64), where Zn from the substrate dissolves into the solution and, hence, supplies the electrons necessary for the reduction, deposition reaction (65). The overall displacement deposition reaction, Eq. (66), is obtained via the combination of the two partial electrode reactions, oxidation and reduction, Eqs (64) and (65), respectively. Thus, in the displacement deposition of Cu on a Zn substrate, a layer of metallic Cu is deposited on the zinc, while Zn dissolves into solution (Fig. 25). We stated that this reaction is possible because the Zn/Zn^{2+} system has a lower electrode potential than the Cu/Cu^{2+} one (Fig. 24).

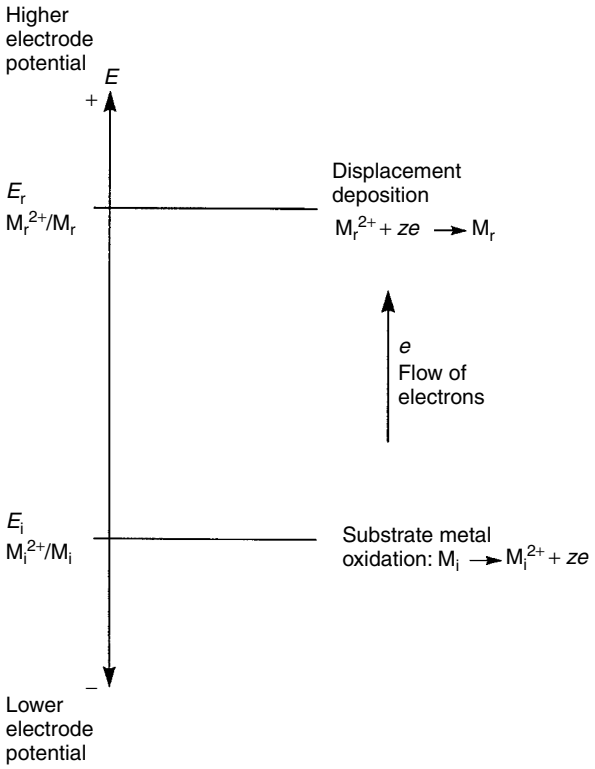


Fig. 26 Relationship between partial reactions in displacement deposition.

The overall displacement deposition reaction according to Eq. (66) can be considered as the reaction of the electrochemical cell

$$Zn/ZnSO_4(aq)//CuSO_4(aq)/Cu \quad (71)$$

3.4.2.1 Predictions of the Feasibility of the Reaction. Complexed Metal Ions in Displacement Deposition

Let us determine whether we can use the displacement deposition technique to deposit Sn on a Cu substrate

$$Sn^{2+} + Cu \longrightarrow Sn + Cu^{2+} \quad (72)$$

The simplest way to determine this is to use the principle presented in Figs 24 and 26. According to this principle, Sn cannot be deposited by displacement on a Cu substrate since the standard electrode potential of the Cu^{2+}/Cu couple is more positive than that of the Sn^{2+}/Sn couple

$$E^{\circ}(Sn^{2+}/Sn) = -0.136\ V \quad (73)$$

$$E^{\circ}(Cu^{2+}/Cu) = +0.34\ V \quad (74)$$

However, reaction (72) can be made to proceed from left to right spontaneously if the potential of the Cu^{2+}/Cu electrode is made more negative than that of Sn^{2+}/Sn electrode. This can be achieved by complexing the Cu^{2+} ions in solution. The preferred complexing agent is the CN^- ion. When CN^- ions are added to the solution of $CuSO_4$, the concentration of Cu^{2+}

ions is reduced and the electrode potential of the Cu^{2+}/Cu electrode is moved to the negative value. We illustrate this in the following example.

If sufficient NaCN (or KCN) is added to the solution of Cu(II) ions to form the complexed ions $[Cu(CN)_3]^{2-}$ and if the excess of CN^- ions is such that the concentration of free CN^- is 1×10^{-4} M, then the concentration of free Cu^+ ions can be calculated from the dissociation constant of the complex. The dissociation constant of the complexed ion $[Cu(CN)_3]^{2-}$ is 5.6×10^{-28}, that is,

$$\frac{(Cu^+)(CN^-)^3}{([Cu(CN)_3]^{2-})} = 5.6 \times 10^{-28} \qquad (75)$$

For 0.05 M $[Cu(CN)_3]^{2-}$ and 1×10^{-4} M CN^- from Eq. (75), one concludes that the concentration of Cu^+ is 2.8×10^{-17} M. The reversible electrode potential of Cu^+/Cu electrode can be calculated for this concentration of Cu^+ from the Nernst equation

$$E = 0.55 + 0.059(\log 2.8 \times 10^{-17}) = -0.43 \text{ V} \qquad (76)$$

where 0.55 is the standard electrode potential for Cu^+/Cu. Thus, in this case Cu^+/Cu is more negative than Sn^{2+}/Sn and reaction (72) will go spontaneously from left to right, resulting in the displacement deposition of Sn on Cu. In this approximate calculation, we considered that only $[Cu(CN)_3]^{2-}$ complexed ions are present in the solution and that the CN^- ions do not affect the electrode potential of the Sn^{2+}/Sn electrode. However, in reality, there is a mixture of different CN^- complexes of copper in the solution.

Reaction (72) is a very important displacement deposition reaction in the printed circuit industry. It is used to help the soldering capability of copper.

3.4.3 Kinetics and Mechanism

The kinetics and mechanism of the displacement deposition of Cu on a Zn substrate in alkaline media was studied by Massee and Piron [62]. They determined that at the beginning of the deposition process the rate is controlled by activation. The activation control mechanism changes to diffusion control when the copper covers enough of the Zn surface to facilitate further deposition of copper. This double mechanism can explain the kinetic behavior of the deposition process.

The mechanisms of the crystal-building process of Cu on Fe and Al substrates were studied employing transmission and scanning electron microscopy [63]. These studies showed that a nucleation-coalescence growth mechanism (Sect. 3.2.4) holds for the Cu/Fe system and that a displacement deposition of Cu on Fe results in a continuous deposit. A different nucleation and growth model was observed in the case of the Cu/Al system. The displacement deposition of Cu on Al substrate starts with the formation of isolated nuclei and clusters of Cu. This mechanism results in the development of dendritic structures.

The properties of deposits may be controlled by changing the kinetics of the deposition and the mechanism of crystallization. One way to achieve this is by complexing the depositing ions, as stated above.

3.5 Electrochemical Deposition of Alloys

3.5.1 Introduction

Sections 3.1 and 3.3 treat kinetics and mechanism of electrochemical deposition

from solutions containing one kind of metal ions. In this case, there is only one metal deposition process (Eq. 1). Section 3.5 discusses the electrochemical deposition from solutions containing two or more kinds of metal ions. This case involves simultaneous processes of the deposition of two or more metals. The result of these simultaneous processes is the formation of an alloy. The alloy-phase structure depends on the electrochemical deposition conditions [64].

Section 3.5.2 discusses electrodeposition and Sect. 3.5.3 discusses electroless deposition.

3.5.2 Electrodeposition

3.5.2.1 Feasibility

The feasibility of the simultaneous deposition of two or more metals at the cathode is solved on the basis of current–potential curves for individual, single (pure) metals. Figure 27 shows current–potential curves for a general case of deposition of single metals M_1 and M_2. It can be seen that if electrodes are immersed in the solution of both M_1 and M_2 ions, and if the cathode potential is at $E(i)_1$ value, only metal M_1 will deposit from the solution. If the cathode potential is at $E(i)_2$ value, both metals will be deposited at the cathode, forming an alloy. From Figure 27 it follows that simultaneous deposition of two metals is feasible if their deposition, or equilibrium (rest), potentials E_{eq} are close.

3.5.2.2 A Simple Model

A simple model assumes that the composition of the alloy at a given potential is determined by the ratio of the values of current densities of deposition of individual metals at that potential; for example, in Fig.27 the composition of the M_1–M_2 alloy deposited at the potential $E(i)$ will be given by the ratio i_1/i_2

$$\frac{M_1}{M_2} = \frac{i_1}{i_2} \tag{77}$$

Using equation for $i = f(\eta)$, Eq. (16), and assuming $a_1 = a_2$, we get

$$\frac{i_1}{i_2} = \left(\frac{i_{o,1}}{i_{o,2}}\right) \exp[af(E_{eq,1} - E_{eq,2})] \tag{78}$$

where $f = F/RT$. Thus, the ratio i_1/i_2 depends on (1) exchange current densities, i_o, and (2) the difference of equilibrium

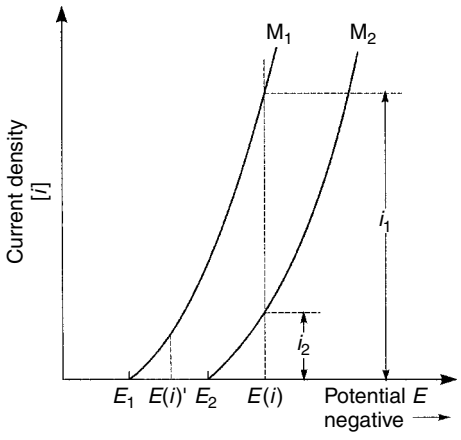


Fig. 27 The simultaneous deposition of metals M_1 and M_2 in which the rest (equilibrium) potential $E_{R,1}$ is more positive than $E_{R,2}$ and $a_1 = a_2$.

potentials, $E_{eq,1} - E_{eq,2}$. Equation (78) shows that the larger the value of $i_{o,1}$ compared with $i_{o,2}$, the greater the amount of metal M_1 deposited compared with the amount of metal M_2 deposited.

Equation (78) assumes that the deposition reactions are activation controlled, $\eta = \eta_{ct}$ (see Sect. 3.1.3). Figure 28 shows current–potential curves for the case when the deposition of metal M_1 is limited by transport ($i_1 = i_{L,1}$; diffusion control). It can be seen that in this case, at the cathode potential $E(i)$, metal M_2 with a more negative equilibrium potential ($E_{eq,2} < E_{eq,1}$) is deposited to a greater extent than metal M_1, $i_2 > i_1$.

3.5.2.3 Extended Feasibility

It is possible to extend the feasibility of codeposition of two metals to cases where this does not look possible owing to a very large difference in the equilibrium potentials of individual metals, Δ ($E_{eq,1} - E_{eq,2}$) (Fig. 27).

One very important example is the electrodeposition of Cu–Zn alloys. Standard electrode potentials for Cu^{2+}/Cu and Zn^{2+}/Zn are 0.3419 and −0.7618 V, respectively (Fig. 29a). The difference in the standard electrode potentials is 1.1037 V. This difference in potentials can be reduced by complexing Cu^{2+} and Zn^{2+} ions in the solution. The preferred complexing agent is the CN^- ion. In a solution containing CN^- ions (NaCN, KaCN), the concentration of Cu^{2+} is reduced much more than that of Zn^{2+} because of the difference in the dissociation constant of the cyanide complex. The dissociation constant of the $[Cu(CN)_3]^{2-}$ complex is

$$\frac{(Cu^+)(CN^-)^3}{([CuCN)_3]^{2-})} = 5.6 \times 10^{-28} \qquad (79)$$

and the dissociation constant of the $[Zn(CN)_4]^{2-}$ complex is

$$\frac{(Zn^{2+})(CN^-)^4}{[Zn(CN)_4]^{2-}} = 1.3 \times 10^{-17} \qquad (80)$$

For 0.05 M $[Cu(CN)_3]^{2-}$ and 1×10^{-4} M CN^-, we calculate the concentration of Cu^+ from Eq. (79) as 2.8×10^{-17} M. The equilibrium electrode potential of Cu^+/Cu electrode, for this concentration of Cu^+ is −0.43 V (see Sect. 3.4). The calculated concentration of Zn^{2+} in 0.025 M $[Zn(CN)_4]^{2-}$, using Eq. (80), is 6.62×10^{-5} and the corresponding equilibrium electrode potential of Zn^{2+}/Zn electrode is −0.85 V (Fig. 29b). Thus, the results of the complexation are as

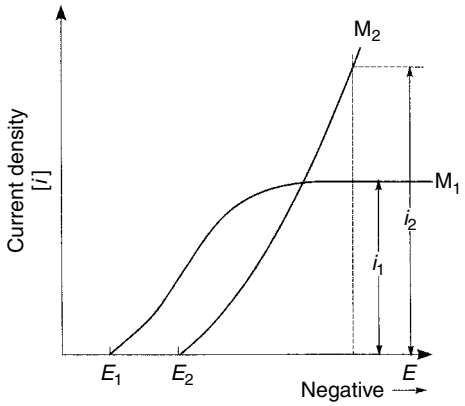


Fig. 28 The simultaneous deposition of metals M_1 and M_2 in which the deposition of metal M_1 is limited by diffusion control; metal M_2 with a more negative rest potential, $E_{R,2} < E_{R,1}$, is deposited to a greater extent than metal M_1, $i_2 > i_1$.

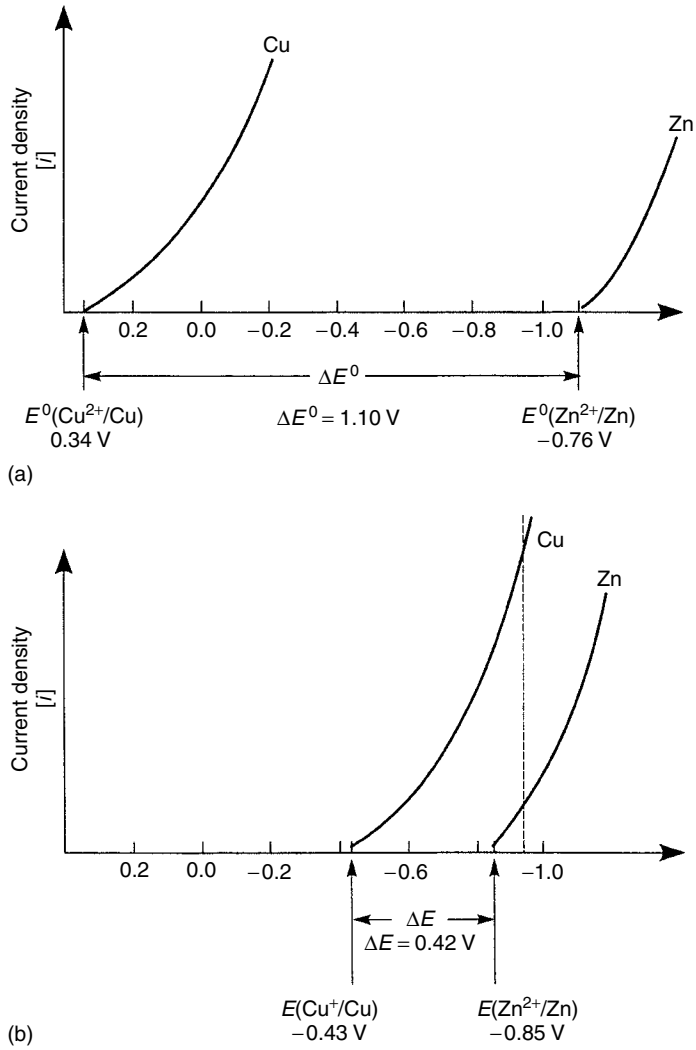


Fig. 29 Feasibility of electrodeposition of Cu−Zn alloys (a) in the absence of complexing agent and (b) in the presence of complexing agent.

follows: (1) the concentration of Cu ions in the solution is reduced much more than that of Zn ions owing to the difference in the dissociation constants of complexes and (2) the electrode potentials of Cu^{+}/Cu are shifted closer to the Zn^{2+}/Zn equilibrium potential so that the difference in equilibrium potentials is reduced from 1.1037 to 0.40 V and under these conditions the electrodeposition of Cu−Zn alloys is feasible (Fig. 29b).

3.5.2.4 Anomalous Codeposition

In the normal deposition of alloys, the more positive (noble) metal deposits preferentially. The result of this is that

the ratio of the more positive metal (M_1) to the less positive metal (M_2) in the alloy deposit is larger than that in the solution. Referring to Fig. 27 and Eq. (77), for the normal deposition we have

$$\left(\frac{M_1}{M_2}\right)_{\text{deposit}} > \left(\frac{M_1}{M_2}\right)_{\text{solution}} \tag{81}$$

In anomalous deposition of alloys the opposite equation is valid

$$\left(\frac{M_1}{M_2}\right)_{\text{deposit}} < \left(\frac{M_1}{M_2}\right)_{\text{solution}} \tag{82}$$

since in this case the less positive (more negative) metal (M_2) deposits preferentially.

A very interesting and technologically important case of anomalous deposition is the electrodeposition of Permalloy, Fe−Ni alloy (see Sect. 3.8). Glasstone and Symes [65] found that in the deposition of Fe−Ni alloys from solutions containing more than about 20% of the total metal as Fe, the initial deposit contains relatively more Ni than the electrolyte.

Most theories of anomalous deposition in this case are based on the studies on kinetics and mechanism of electrodeposition of iron by Bockris et al. [66]. They proposed the following mechanism for the electrodeposition of Fe

$$Fe^{2+} + H_2O \longleftrightarrow [FeOH]^+ + H^+ \tag{83}$$

$$[FeOH]^+ + e \xrightarrow{\text{RDS}} FeOH \tag{84}$$

$$FeOH + H^+ + e \longleftrightarrow Fe + H_2O \tag{85}$$

where RDS stands for rate-determining step. The overall electrode reaction is

$$Fe^{2+} + 2e \longrightarrow Fe \tag{86}$$

According to this mechanism, the electrochemically active species are $[FeOH]^+$ and FeOH. Iron-hydroxide ions result from the hydrolysis of Fe^{2+} ions in the solution, Eq. (83).

Matulis et al. [67] have suggested a similar mechanism for the electrodeposition of Ni involving the hydrolysis of Ni^{2+} ions and the formation of nickel-hydroxide ions, $[NiOH]^+$. Nickel-hydroxide ions are the electrochemically active species in the electrodeposition of Ni.

Thus, hydrolysis plays a significant role in the electrodeposition of the single metals, Fe and Ni.

During the deposition of iron and nickel, H^+ (H_3O^+) discharge occurs and the pH *at the electrode surface* increases. Dahms and Croll [68] derived an equation to evaluate the surface pH as a function of the bulk pH. Haris [69] calculated the change in surface pH during the electrodeposition of a bivalent metal.

Hessami and Tobias [70] extended the mechanisms of Bockris et al. and Matulis et al. of the deposition of single metals (Ni, Fe) to the mathematical modeling of codeposition of Ni−Fe alloys. This mathematical model for the anomalous alloy deposition describes the electrode processes using the calculated interfacial concentrations. The inhibition (reduction) of nickel partial current density during alloy deposition and the anomalous deposition are explained on the basis of the relative concentrations of metal-hydroxide ions, $[MOH]^+$. Calculations show that the $[FeOH]^+$ concentrations are higher than $[NiOH]^+$ because it has a much smaller dissociation constant ($[MOH]^+ = (M^{2+})(OH^-)/K_{\text{diss}}$). $[FeOH]^+$ ions compete with $[NiOH]^+$ ions for surface sites, and the result of this competition is inhibition (decrease) in Ni deposition in the presence of $[FeOH]^+$ ions. Figure 30

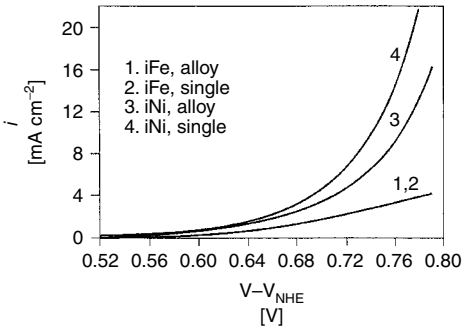


Fig. 30 Calculated partial currents *I* of nickel and iron, single versus alloy at 1600 rpm (from Ref. 70 with permission from the Electrochemical Society).

shows calculated partial current densities of Ni and Fe, single versus alloy deposition. Thus, the model of Hessami and Tobias explains the basic features of anomalous deposition: (1) the ratio of the less positive (noble) metal, Fe, is much higher in the alloy than in the electrolyte, and (2) the presence of the less positive (noble) metal inhibits the discharge of the more positive (noble) metal, Ni.

The model based on metal-hydroxide ions ($[MOH]^+$) was further developed by Grande and Talbot [71]. Sasaki and Talbot [72] demonstrated the extendibility of this model to the electrodeposition of Co–Fe and Ni–Co alloys. They found that there is a slight inhibition of the more positive metal deposition and a promotion (acceleration) of the less positive metal deposition for all binary iron-group alloys.

Golodnitsky et al. found that inhibition of the more positive metal deposition by the less noble one does not depend on the anion composition of the electrolyte [73].

3.5.3 Electroless Deposition

Electroless (autocatalytic) deposition takes place on *catalytic surfaces* (see Sect. 3.3 and Chapter 4 in Volume 2). The anodic partial reaction in electroless deposition is strongly dependent on the catalytic properties of the initial substrate (e.g. Pd catalytic nuclei) and the deposited metal (e.g. Cu in the electroless deposition of Cu). In the electroless deposition of alloys, the catalytic properties of all the components of an alloy are important. Ohno et al. [74] evaluated the catalytic activity of metals for the anodic oxidation of different reducing agents and obtained the catalytic activity series on the basis of the potentials at a reference current density. This series may be used to choose the reducing agent suitable to the metal, or alloy, to be deposited.

Cobalt- and nickel-based alloys may have a variety of applications in modern silicon and magnetic recording technologies; for example, Co–Ni(P) alloys can be used for the fabrication of electrical contacts [75], Co–W(P) can be used for the fabrication of diffusion barriers in semiconductor technology [76], Ni–W(P) can be used as heating resistors [77], and a variety of binary and ternary Co alloys can be used for the fabrication of magnetic recording media [78]. A complete review of the electroless deposition of alloys is given by Ohno [79].

3.6 Structure and Properties of Deposits

3.6.1 Introduction

In considering the structure of metals, we distinguish three basic structures: (1) atomic bulk and surface structures, (2) electronic structure of metals and surfaces, and (3) microstructure.

From atomistic aspects, a metal can be considered as a fixed lattice of positive ions permeated by a gas of free electrons [1]. Positive ions are the atomic cores, while the electrons are the valence electrons. Since there are about 10^{22} atoms in 1 cm^3 of a metal, one can expect that some atoms are not exactly in their right place. Thus, one can expect that a real lattice will contain defects (imperfections). The most common defects are point defects (e.g. a vacancy, an interstitial) and dislocations (e.g. the edge dislocation, screw dislocation) [2].

The electronic structure of metals and metallic surfaces are interpreted by the free-electron theory of metals [1, 2].

The volume over which a fixed lattice of positive ions extends uniformly is called a *single crystal*. Most metal deposits are not single crystals, but are agglomerates of small crystallites packed together. These crystallites are called *grains*. The boundaries between these grains are lattice defects and are called *grain boundaries*. A microstructure is defined by grain size, grain shape, and grain orientation. The major grain structures are columnar, equiaxed, dendritic, and whisker.

Since there is a strong interrelationship between the structure and the properties of deposits, this section discusses both these subjects.

In Sect. 3.7.1, we give an example of how structure parameters determine the properties of thin metal films and in Sect. 3.7.3, we discuss the influence of microstructures of the substrate on the structure and properties of electrodeposits.

3.6.2 Overpotential Dependence of the Growth Forms

The dependence of growth forms on overpotentials originates from the potential dependence of nucleation and growth processes. Competition between nucleation and growth processes is strongly influenced by the potential of the cathode. Thus, major factors determining growth forms are as follows: (1) Structure of the double layer including concentration of different species present in the solution. Potential dependence of adsorption of additives and their effect on growth forms is discussed in Sect. 3.2.5. (2) Surface structure including concentration of adions at the surface and lattice defects (e.g. dislocations, vacancies). (3) The radius and the size of the critical nucleus (see Sect. 3.2.2). (4) Rate of nucleation (see Sect. 3.2.2). (5) Lateral and vertical growth rates of the crystal grains.

Seiter et al. [80] found a correlation between the overpotential and growth forms of electrodeposited copper on copper sheet substrate with (100) texture, as shown in Fig. 31. Barnes et al. [81] observed similar results on copper single crystal surfaces near (100) face: below 10-mV ridges, 40–70-mV platelets, 70–100-mV blocks and fine platelets and above 100-mV polycrystalline deposit.

3.6.3 Columnar Microstructure

Columnar microstructure, perpendicular to the substrate surface, is shown schematically in Fig. 32. This microstructure is

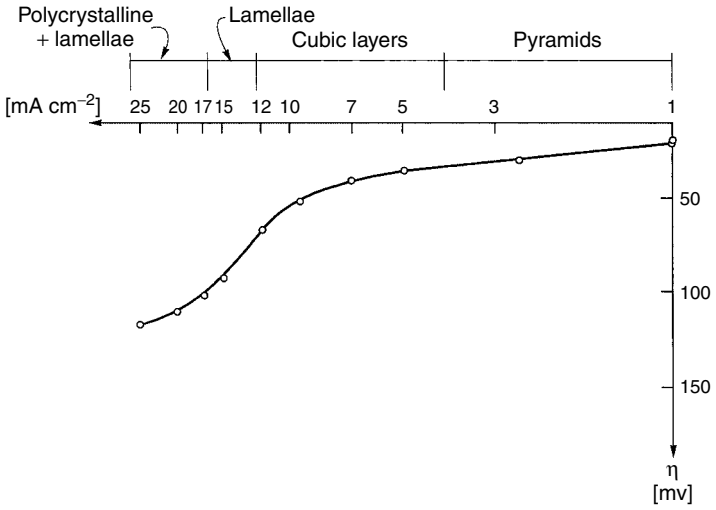


Fig. 31 Current–potential curve showing correlation between overpotential η and growth forms of electrodeposited copper from N $CuSO_4$ and NH_2SO_4, 25 °C (from Ref. 80 with permission from Elsevier Science Inc.).

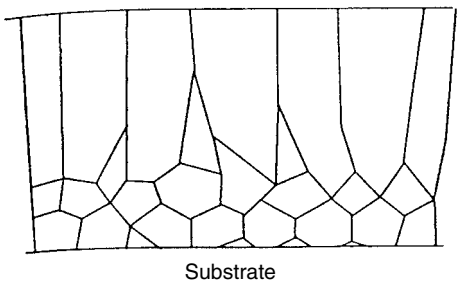


Fig. 32 Schematic cross section (perpendicular to the substrate) of the columnar deposit.

composed of relatively fine equiaxed grains near the substrate which then changes to the columnar microstructure with much more coarse grains at larger distances from the substrate. Development of the columnar microstructure can be interpreted on the basis of *growth competition* between adjacent grains in a similar way as in the development of texture (Sect. 3.6.4). The low–surface energy grains grow faster than the high-energy ones. This rapid growth of the low–surface energy grains at the expense of the high-energy grains results in an increase in the mean grain size with the thickness of deposit and the transition from a fine grain size near the substrate to a coarse columnar grain size. Columnar structures develop in deposits prepared by electrodeposition and evaporation. Srolovitz et al. [82] developed a theoretical model for the growth of columnar microstructures in vapor-deposited films. We find this work interesting for electrodeposition too. It shows that the microstructure of the evaporated film, parallel to the substrate surface, at different depths, has a relatively uniform microstructure composed of fine grains near the substrate

surface and a bimodal grain size distribution at intermediate depths.

Electrochemical deposition of copper at very slow rates leads to a columnar structure with deep crevices [83]. This structure is explained theoretically to be a result of competition between local and nonlocal growth effects.

3.6.4 Development of Texture

Here we examine models that try to explain how textures develop during deposition on oriented (single crystal), textured, polycrystalline, and amorphous substrates. We select electrodeposition of nickel as a model system.

3.6.4.1 Single-crystal Substrate

Single crystals are selected as substrates for two main reasons: (1) to eliminate the effect of grain boundaries and (2) to eliminate the heterogeneity of the substrate due to the presence of grains of various orientations and, thus, various exchange current densities. Single-crystal substrates exert a strong *epitaxial influence* on the growth process. In epitaxial growth, the structure of the deposit is a continuation of the structure of a substrate metal. Froment and coworkers [84] have shown that the epitaxial Ni growth from Watts solution starts with the formation of TECs. TECs then coalesce to give a continuous epitaxial layer. Further growth is complex and involves two different nucleation processes: 3D epitaxial nucleation and independent nucleation. The first process includes epitaxial growth and the second one includes nonepitaxial growth. The balance between these competitive processes depends on the substrate orientation and the deposition conditions; for example, on a (100) Cu face, epitaxial growth continues beyond a Ni deposit thickness of 10 μm. In this case, further growth via the independent nucleation is predominant, and the transition occurs from the epitaxial growth to the polycrystalline growth.

3.6.4.2 Polycrystalline Randomly Oriented Substrate

Electrodeposition on a randomly oriented polycrystalline substrate can result in the development of *preferred orientation*, or *texture*, in thicker deposits. In a polycrystalline material, the crystallographic axis of individual grains (individual crystallites) that constitute the material are randomly oriented with respect to the axes of a fixed reference system (Fig. 33a). If one or more

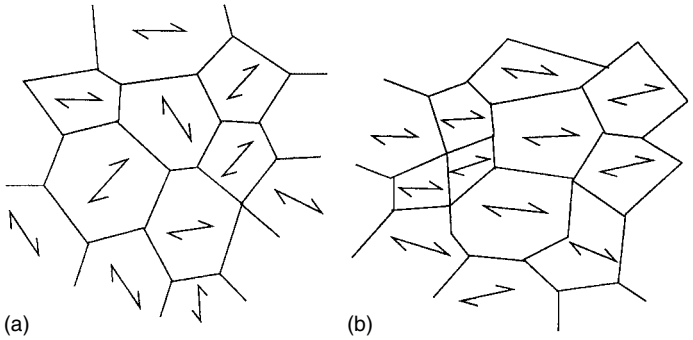


Fig. 33 Schematic representation of polycrystalline randomly oriented substrate (a) and substrate with preferred orientation (texture) (b).

crystallographic axes of grains constituting a polycrystal are fixed (have the same orientation) with respect to the axes of the reference system, the polycrystalline material exhibits preferred orientation or texture (Fig. 33b). The development of texture can occur during the deposition or the postdeposition processing.

The *competitive growth model* of the development of texture during deposition is based on the idea that different crystal faces have different rates of growth. Thus, there is a growth rate competition between crystallites of various orientations. Crystallites of various orientations could be generated either during the *preferential nucleation process* or during the competitive growth mechanism subsequent to the stage of coalescence. The type of texture depends on the composition of electrolyte, substrate, the overpotential, and so on.

3.6.4.3 Amorphous Substrate

An amorphous substrate (e.g. vitreous carbon) is without any epitaxial influence; it is inert with respect to the growth process of the deposit. Froment and coworkers [84] showed that in the initial stages of the Ni growth from Watts' solution on vitreous carbon substrate, orientation of individual 3D nuclei is random. They also showed that a newly coalesced compact deposit has perfectly random orientation. From these observations, they concluded that the texture of thicker Ni deposits in this case is the result of a competitive growth mechanism occurring in a stage of growth subsequent to the coalescence stage.

In Sect. 3.7.1, we present one example which shows how texture may determine the physical properties and reliability of thin films.

3.6.5 Dendrites

Dendritic deposits grow under mass transport–controlled electrodeposition conditions. These conditions involve low concentration of electrolyte and high current density. A dendrite is a skeleton of a monocrystal consisting of stem and branches. The shapes of the dendrites are mainly determined by the directions of preferred growth in the lattice. The simplest dendrites consist of the stem and primary branches. The primary branches may develop secondary and tertiary branches. The angles between the stem and the branches, or between different branches, assume certain definite values in accordance with the space lattice. Thus, dendrites can be two dimensional (2D) or three dimensional (3D).

Figure 34 shows the fractal structure of a flat 2D electrodeposited copper leaves at increasing magnification: 64, 256, 523, and 1024. Fractal objects are the same on different observation scales. The thickness of the deposit in Fig. 34 is of the order of 200 nm. Copper is in its normal fcc structure, but has many microcrystals whose size is not larger than 20 nm.

The dendrite growth process may be described on the basis of *cluster growth model of diffusion-limited aggregation* (*DLA*) and fractal concepts in surface growth [83, 85].

3.6.6 Whiskers

Whiskers grow in the presence of large growth-inhibiting molecules, which usually show a preferential adsorption on certain crystal plains. Incorporation of adsorbed particles in the crystal [86] and internal stress generated during the electrodeposition are major causes of whisker

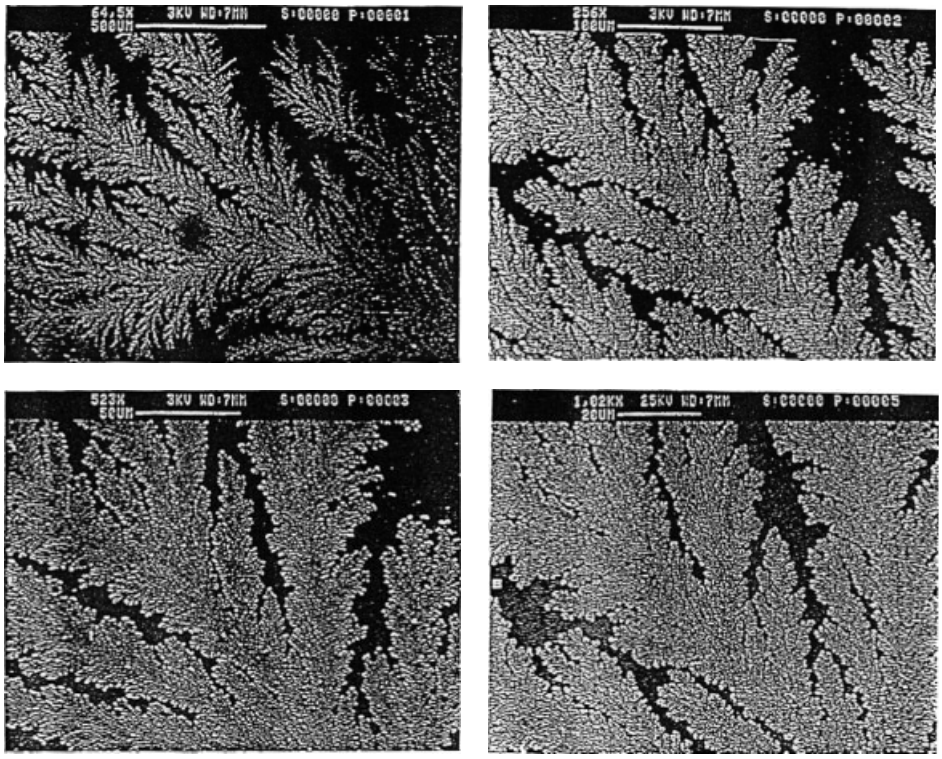


Fig. 34 Sequence of pictures showing different magnifications of the same leaf (from Ref. 85 with permission from Materials Research Society).

formation [87]. The two important characteristics of whiskers are (1) large ratio between the longitudinal and the lateral dimensions and (2) preservation of the lateral dimensions during the growth [30].

3.6.7 Electrical Resistance

According to the modern quantum electronic theory [1, 88, 89], electrical resistivity of a metal results from the scattering of electrons by the lattice. In a perfect lattice, electrons experience no scattering, and they can carry current without any attenuation. A real metal lattice departs from perfect long-range order, and electrons are scattered by imperfections in the lattice potential.

If we assume that all the current is carried by carriers with the same effective mass and velocity, then the resistivity ρ is given by

$$\rho = \Sigma_s \rho_s \tag{87}$$

where ρ_s is the resistivity introduced by the scattering event s. The major components of the lattice imperfection are

$$\rho_l = \rho_{imp} + \rho_{gb} + \rho_{disl} \tag{88}$$

where ρ_{imp} is the impurity resistivity, ρ_{gb} is the grain-boundary resistivity, and ρ_{disl} is the dislocation resistivity.

Knowledge of the electrical resistance as a function of temperature is of great importance in the semiconductor technology (Sect. 3.7) since in the fabrication of integrated circuits (ICs), silicon chips are subjected to multiple curing (annealing) cycles. The significance of understanding the annealing processes is demonstrated here for the case of bilayer Co(P)/Cu. Thin films of electroless Co(P) may be considered as cladding materials for copper conductors in the fabrication of ICs. The concern here is the interdiffusion and, thus, the integrity of copper conductors. Diffusion of Co into Cu conductors could increase the electrical resistance of Cu. Paunovic et al. [90] found that during the heat treatment, the single-phase structure of as-deposited Co(P) is transformed into the two-phase structure of Co and Co_2P. The metallic compound Co_2P is formed at temperatures above 450 °C. The two-phase structure of Co(P) is a crystal mixture of hexagonal-close-packed (hcp) Co and Co_2P. In addition, after heating to 440 °C, hcp Co changes into a mixture of hcp and fcc Co. Since Co_2P is a high-resistivity semiconductive material, the resistivity component ρ_{imp}, Eq. (88), increases after the formation of the new phase. These structural changes can explain the increase of the sheet resistance at the end of the thermal cycle [90]. This example illustrates the significance of understanding the annealing processes in the fabrication of ICs well.

3.6.8 Mechanical Properties

One interesting example is electroless Ni(P). Electroless Ni(P) is harder and has better corrosion resistance than electrodeposited Ni(P). Nonmagnetic electroless Ni(P) is used as an underlayer in high-density metallic memory disk fabrication to improve the mechanical finish of the surface. Thus, hardness, wear resistance, and corrosion resistance have been major properties determining the technological applications of electroless Ni(P) in the electronic, aerospace (stators for jet engines), automotive, machinery, oil and gas production, power generation, printing, and textile industries.

It is interesting to note that Brenner and Riddell [1, 43] ''accidentally encountered'' electroless deposition of nickel and cobalt during the electrodeposition of nickel–tungsten and cobalt–tungsten alloys (in the presence of sodium hypophosphite) on steel tubes in order to produce material with better hardness than steel. They found deposition efficiency to be higher than 100%, which was explained by the contribution of electroless deposition to the electrodeposition.

3.7 Applications in Semiconductor Technology

3.7.1 Introduction

The most active areas in the modern applications of electrochemical deposition are semiconductor technology and magnetic recording.

One major recent advance in silicon-based semiconductor industry is the development of copper interconnects on chips. This new technology replaces the traditional aluminum or aluminum alloy (e.g. Al–Cu) conductors produced by physical vapor deposition (PVD) with copper conductors manufactured by electrodeposition. Copper has been replacing aluminum since 1999 owing to its low bulk electrical

resistivity and superior electromigration resistance in comparison with aluminum. The electrical resistivity of pure Al and Cu are 2.9 and 1.7 μΩcm, respectively, and that for Al alloys is 3–4 μΩcm. The activation energies for electromigration, using identical structures and experimental conditions, are 0.81 and 1.1 eV for Al (0.5 wt% Cu) and Cu, respectively [91–94]. Lower electrical resistivity results in higher speed of devices. Higher electromigration resistance (higher activation energy for electromigration) results in higher reliability and, thus, lower failure rate of interconnects. The related measure of reliability, the median time to failure (MTF) t_{50} (time to 50% cumulative failure), due to electromigration is given by an Arrhenius-type equation

$$t_{50} = AJ^{-n}\exp(E_a/kT) \qquad (89)$$

where E_a is the activation energy for grain-boundary diffusion, n is the current density exponent, J is the current density, T is the temperature, k is the Boltzmann's constant, and A is a constant [92, 95].

The number of transistors per chip is increasing continuously (for a microprocessor, the total number of transistors per chip was 11, 21, and 40 M (million) for years 1997, 1999, and 2001, respectively) and the physical feature size of transistors is decreasing; consequently the dimensions of interconnections (interconnects) on the chip are scaling down; for example, linewidths were 0.25, 0.18, 0.15, and 0.10 μm in the years 1997, 1999, 2001, and 2006, respectively. This scaling down of interconnects on chips requires interconnect metal of high quality and better understanding of physical properties of thin films [96].

The required degree of understanding of physical properties of thin metal films used for interconnects on chips is illustrated in the following example. It was found that the performance of conductors on chips, Al or Cu, depends on the structure of the conductor metal; for example, Vaidya and Simha [97] reported that the measured MTF of Al-0.5% Cu thin films is a function of three microstructural variables (attributes): its median grain size s, the statistical variance σ^2 of the grain size distribution, and the degree of the (111) fiber texture in the film. They showed that the dependence of electromigration resistance, R_{em}, on microstructure can be expressed as the product of the median grain size and normalized film texture, when divided by the variance of the grain size spread

$$R_{em} = k(s/\sigma^2)\log(I_{111}/I_{200})^3 \qquad (90)$$

where k is a constant. This equation shows that strong (111) texture is beneficial for the reliability of Al−Cu interconnects. It was proposed that the above equation could be used to predict the electromigration resistance of Al thin films, independent of the deposition technique. The corresponding equation for Cu interconnects has not yet been established. However, there is extensive work in progress in this direction.

In this section, we review the present manufacturing processes and discuss the challenges facing the extendability of these processes to the new generation of IC products. There are three primary fabrication processes: (1) metal-reactive ion etching (RIE), (2) dielectric RIE, or damascene, and (3) through-mask deposition process.

The metal-RIE process is used in the fabrication of Al interconnects on chips. In this process, a blanket thin film of Al (or Al alloys, like Al−Cu, Al−Si) is deposited and then etched in a reactive plasma (RIE) through a photoresist stencil. After RIE, a

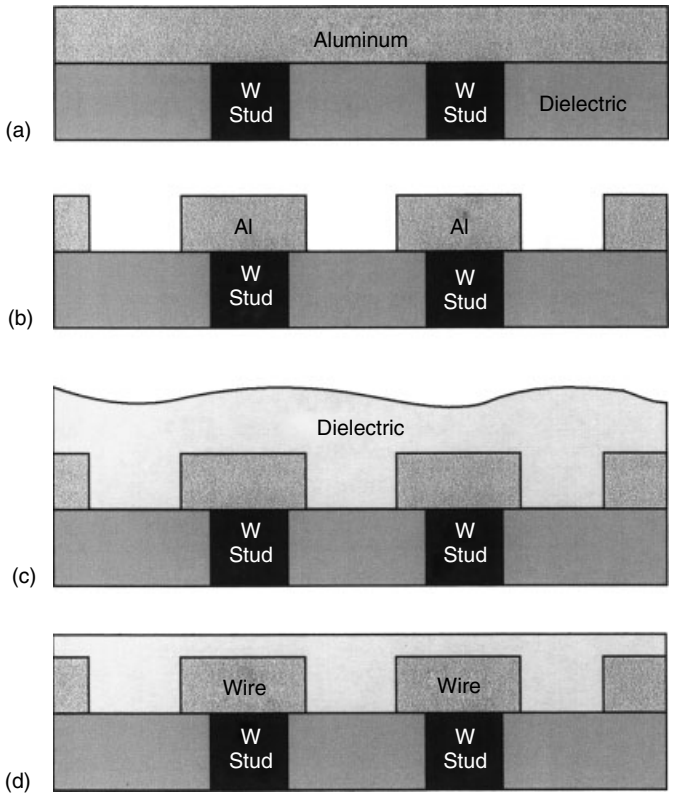


Fig. 35 Process steps for forming Al interconnects using metal-RIE process (a) Al deposition, (b) Al RIE through photoresist stencil, (c) dielectric (e.g. SiO_2) deposition, (d) dielectric CMP.

dielectric is then deposited so that it fills the gaps between the lines as well above them (Fig. 35).

3.7.2 Deposition of Cu Interconnections on Chips

The change from Al to Cu interconnects required a change in the fabrication process from metal RIE to dielectric RIE since it is difficult to pattern Cu by RIE [98].

Multilevel Cu interconnections on chips are now fabricated using a dielectric-RIE process. In this process, a blanket Cu deposition is followed by chemical–mechanical polishing (CMP) of Cu [99]. This approach is known as a *damascene process* (process that has been used in Damascus for centuries to form inlaid metal features on jewelry). Figure 36 illustrates the process sequence for a single damascene process. This IC fabrication process starts with the deposition of a dielectric layer on Si wafer, patterning it using photolithography and dielectric RIE. After patterning the dielectric, the barrier metal and the Cu seed layer are deposited using physical or chemical vapor deposition (CVD). Finally, Cu is electrodeposited into the recesses, trenches (lines) and vias (holes). The excess of the deposited Cu on the upper surface is removed by CMP [99]. In the

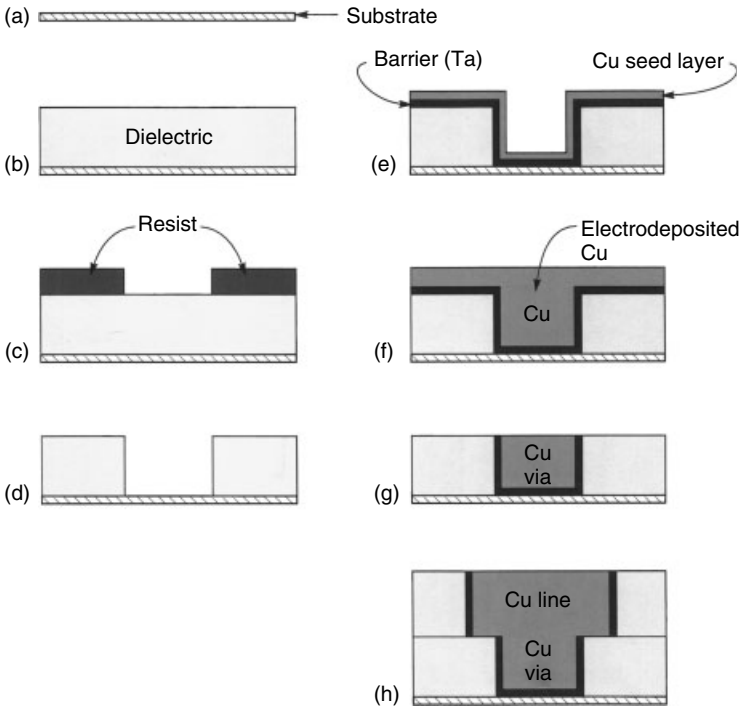


Fig. 36 Process steps for forming Cu interconnects using single damascene process (dielectric patterning); (a) substrate, (b) dielectric deposition, (c) dielectric RIE through photoresist stencil, (e) PVD or CVD of diffusion barrier and Cu seed layer, (f) electrodeposition of Cu into via (vertical interconnection), (g) CMP of Cu excess, (h) patterning and deposition of Cu line (wire).

dual damascene process, vias and trenches are patterned and filled with Cu at the same time (Fig. 37).

Selective Cu deposition process was suggested as an alternative process for the fabrication of multilevel Cu interconnections (Fig. 38). This process is known as *through-mask deposition* process [100]. The first step in this process is the deposition of a Cu seed layer on Si wafer; then a resist mask is deposited and patterned to expose the underlying seed layer in vias and trenches. In the next step, Cu is deposited to fill the pattern. After the Cu deposition mask is removed, the surrounding seed layer is etched, and the dielectric is deposited. Electroless Cu deposition was suggested for blanket and selective deposition processes [101, 160].

A variation of the selective Cu deposition process, limited to electroless Cu deposition, is the *lift-off process*, known as a *planarized metallization process* [102]. Figure 39 shows a process sequence for this technology.

There is a basic difference between the damascene and the plating-through-mask processes in the way in which the trenches and vias are filled with electrochemically deposited Cu, either through electro- or electroless technique. In multilevel metal structures, vias provide

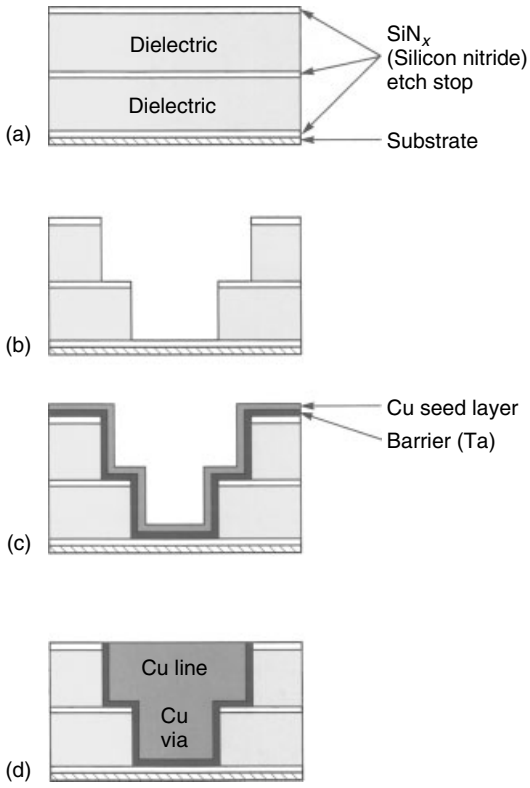


Fig. 37 Process steps for dual damascene process; (a) deposition of dielectric, (b) dielectric RIE to define via and line, (c) deposition of diffusion barrier and Cu seed layer, (d) electrodeposition of Cu into via and trenches followed by Cu CMP.

a path for connecting two conductive regions separated by interlevel dielectric (ILD). In a damascene process, the Cu deposit grows from the active bottom and sidewalls (Fig. 40a). Preferred growth from the bottom may be achieved by the addition of suitable additives [103, 104]. In the plating-through-mask process, the bottom is active, the sidewalls are inactive, and Cu deposit grows from the bottom, Fig. 40b [105]. In all these considerations, one has to take into account the active surface area and current distribution [106, 107].

One of challenges in the fabrication of interconnections on chip is the electrodeposition of Cu in vias of small diameter ($<0.2\,\mu m$) and high aspect ratio, height/width $\gg 2$. Modeling of these processes shows that new Cu electrodeposition solutions and new deposition techniques are necessary to solve problems introduced by the development of new ICs [108].

3.7.3 Diffusion Barriers and Seed Layer

Copper introduces new problems in the fabrication of interconnects on chips. The most important one is the diffusion of Cu into Si, SiO_2, and other dielectrics [92], and reaction of Cu with Si forming silicides [109]. Diffusion of Cu through Si results in the poisoning of devices (transistors) and diffusion through SiO_2, dielectrics, leads to the degradation of dielectrics, which can result in

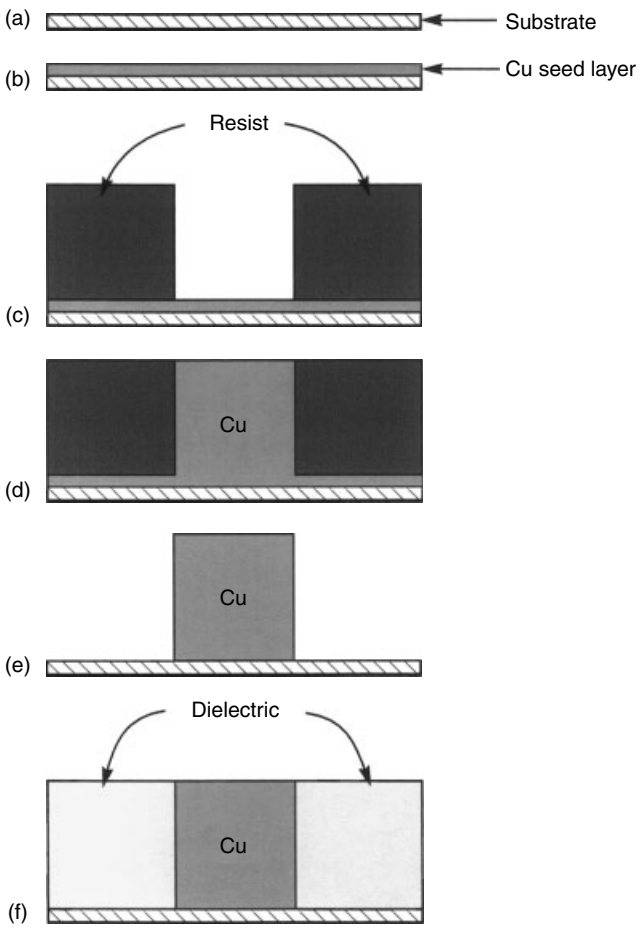


Fig. 38 Through-mask deposition process. (a) Substrate, (b) PVD or CVD Cu seed layer deposition, (c) photoresist deposition and patterning, (d) through-mask electroless deposition of Cu, (e) stripping of photoresist and etching of Cu seed layer outside of line, (f) dielectric deposition.

shorts between interconnects. Thus, diffusion barrier layers are an integral part of the fabrication of copper interconnects (Figs 36, 37). Barrier films isolate (encapsulate) Cu interconnects from adjacent dielectric material. The most studied diffusion barriers are Ta [110, 111], Ti, and TiN [112, 113].

Since barrier metals have relatively high electrical resistivity (Ta 12.4; Ti 80 μΩcm) it is necessary to cover the barrier layer with a conductive metal layer (seed layer). This conductive metal layer is a *Cu seed layer*, which is deposited using PVD or CVD techniques (Figs 36, 37). When the electrodeposition of Cu on a barrier/Cu seed layer bilayer is finished, vias and trenches are filled with Cu, and excess of Cu is removed using CMP (Figs 36g,h, 38d). The exposed Cu in lines needs *capping* with a

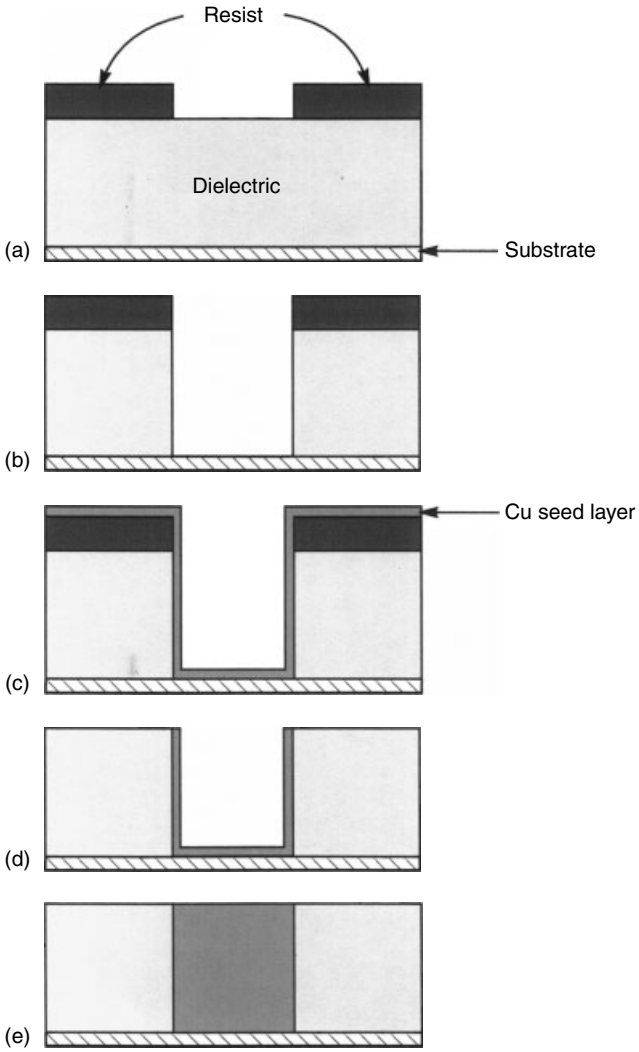


Fig. 39 The process sequence for the lift-off process (the planarized metallization process); (a) a resist film is patterned on a dielectric film, (b) dielectric patterning, (c) thin catalytic film layer (PVD or CVD Ti, Al) is deposited, (d) a lift-off technique removes the excess material, leaving the catalytic layer in the trench only, (e) electroless Cu deposition.

barrier material to prevent diffusion. Two types of capping material may be used: metals (alloys) and dielectrics. The suggested selectively deposited metal barriers (Fig. 41a) are electroless Co(P), Co−Ni(P), Co−W(P), and selective CVD(W) [114, 115]. The preferred dielectric barrier material (Fig. 41b) is blanket SiN_x. The result of capping is a fully encapsulated Cu line (Fig. 41).

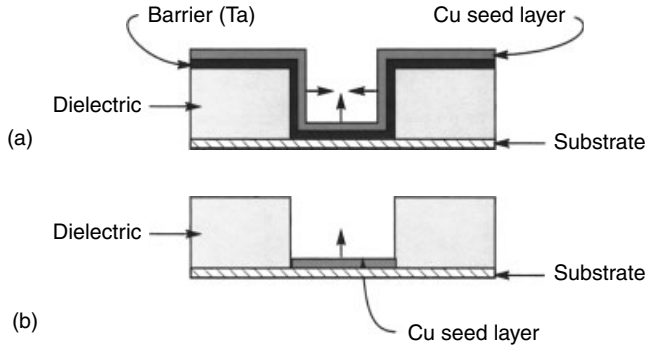


Fig. 40 Growth of deposit in vias and trenches during Cu electrodeposition in (a) damascene and (b) deposition-through-mask process.

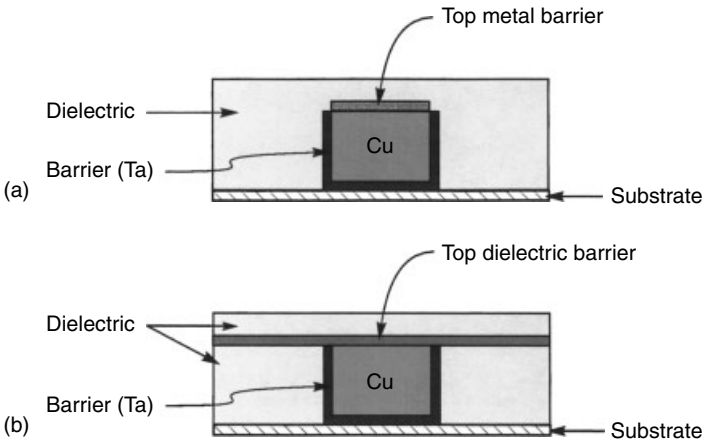


Fig. 41 Fully encapsulated Cu line; (a) metal (alloy) top barrier and (b) dielectric top barrier.

It can be seen from the above discussion that Cu is electrodeposited in vias and trenches on a bilayer: barrier metal/Cu seed layer. In cases when the barrier layer is composed of two layers (e.g. TiN/Ti), Cu is electrodeposited on a trilayer: barrier bilayer/Cu seed layer. This type of underlayer for the electrodeposition of Cu raises a series of interesting theoretical and practical questions of great significance regarding the reliability of interconnects on chips. In Sect. 3.7.1, we have shown that the reliability of interconnects depends on the microstructural attributes of the electrodeposited Cu (for Cu-based interconnects). These microstructural attributes, such as grain size, grain size distribution, defect density, and texture, determine the mechanical and physical properties of thin films. Thus, one basic question in the above-mentioned series of questions is the problem of the influence of an underlayer barrier metal on the microstructure of the Cu seed layer. The second question

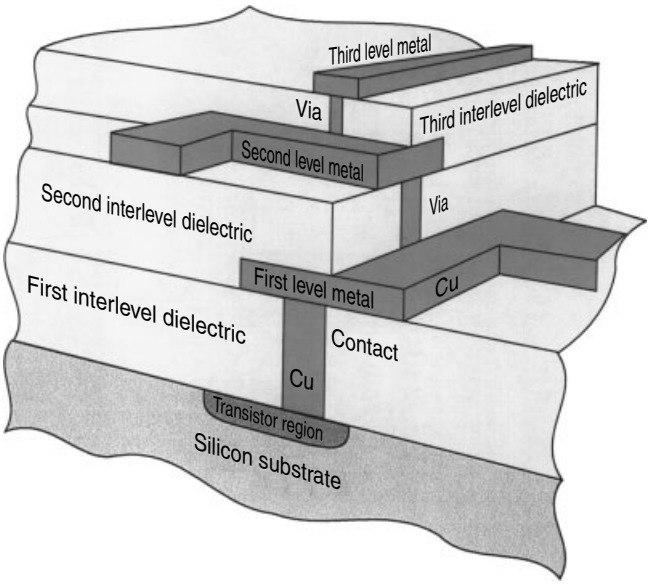


Fig. 42 Three-level Cu interconnect IC structure.

is the influence of the microstructure of Cu seed layer on the structure and properties of electrodeposited Cu [116]. Zielinski et al. [116] showed that Cu microstructure is sensitive to the texture and microstructure of the barrier metal; for example, textural inheritance was observed on Ta underlayer. In this case, there is quasi-grain-to-grain epitaxy. In addition, the presence of a barrier layer can further influence microstructural evolution upon annealing. Tracy et al. [117] determined that a strongly textured underlayer, such as Ti or Ti/TiN, induces a similar strong texture in the Cu. A strong Ti texture is passed to TiN, which passes the texture to Cu (or to Al—Cu). Tracy and Knorr [118] determined that a typical copper film consists of three texture components: (111), (200), and random. Indeed, (200) and (111) texture components are possible under some deposition conditions.

Integration of Cu with a dielectric introduces new problems and challenges [119] (Fig. 42); for example, if polyimides are used as intermetal dielectrics, the reliability concerns are the corrosion of the underlying metal and the adhesion of metal films to the polyimide underlayer (101).

The above discussion on the influence of underlayer on microstructure and reliability of interconnects in IC illustrates the need for an in-depth understanding of the process of deposition and the physical and mechanical properties of electrodeposited Cu used in IC fabrication. There are great challenges and great opportunities for high-level studies of electrochemical deposition.

3.8 Applications in Magnetic Recording Industry

3.8.1 Introduction

In the present technology, digital magnetic recording system in the computer industry

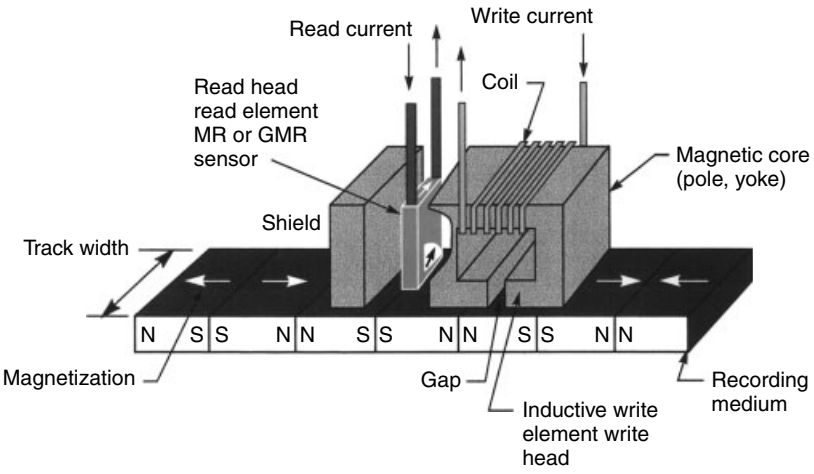


Fig. 43 Schematic representation of digital magnetic recording system consisting of (1) inductive magnetic write head, (2) shielded MR (magnetoresistive) or GMR (giant magnetoresistive) head, and (3) magnetic recording medium (from Ref. 120 with permission from IBM Research and Development Division).

consists of three components: inductive magnetic write head, giant magnetoresistive (GMR) read head, and magnetic recording medium, Fig. 43.

Electrochemical deposition is applied in the fabrication of all the three components in at least one layer, magnetic or nonmagnetic. Magnetic layers and Cu coil in the write head electromagnet are fabricated by electrodeposition. Magnetic layers in the read head and in the recording medium are fabricated by vacuum deposition techniques. Electrodeposition is used for the fabrication of shields in the GMR read head, and electroless deposition is used to fabricate the underlayer in the rigid (also called *hard*) disk magnetic medium. In this section, we discuss all the three components of a magnetic recording system.

3.8.2 Write Heads

Inductive magnetic write (recording) head consists of three main structural parts: C-shaped magnetic core (pole, yoke), gap, and coil (Fig. 43). It can be seen that a write head is an electromagnet. Electrical current in the coil generates a magnetic field across the gap of the write head. This magnetic field magnetizes the moving magnetic recording medium below. The writing current has one magnitude and two possible directions (+ or −), corresponding to positive and negative saturation magnetization of the storage medium. Thus, this type of recording is digital, that is, there are only two states of magnetization: one state for 1 bit and the other state for 0 bit. These digital magnetic recordings (data) on the recording surface are discrete (quantized), magnetized domains. The storage medium must be capable of retaining the sequence of permanent magnetic states. This is related to the hysteretic behavior of the magnetic recording medium (Appendix A1).

Most (thin-film) inductive magnetic heads have a magnetic core and coil fabricated by electrodeposition using

the plating-through-mask technique (see Sect. 3.7.2). The substrate is usually made of a ceramic, such as Al_2O_3TiC (aluminum titanium carbide) about 2-mm thick. The bottom pole is electrodeposited first. Poles are made of magnetically soft (Appendix A2) metallic alloy, the most common being Permalloy, an alloy of 80 at% Ni and 20 at% Fe (henceforth referred to as $Ni_{80}Fe_{20}$). Since electroplating cannot be done on an insulating ceramic substrate, a seed layer of Permalloy of thickness 80–100 nm is first deposited on the substrate by sputtering. The Permalloy poles are single layered and about 2–4-μm thick. After the bottom pole deposition, a layer of an insulator, SiO_2 or an organic polymer, is deposited.

After the bottom pole and insulator, a microwinding Cu coil is electrodeposited [121]. The insulator has to be prepared for the electrodeposition of Cu. This preparation involves the deposition of Cr/Cu bilayer by sputtering or evaporation. First, a thin layer (10 nm) of Cr is deposited onto the insulator. The function of the Cr layer is to provide a bonding layer between the insulator and Cu. A thin (50–100 nm) layer of Cu seed layer is then sputter deposited on Cr layer to provide sufficient electrical conductivity for subsequent electrodeposition of Cu. Cu is electrodeposited using deposition-through-mask technique. After electrodeposition of Cu coil, an insulator layer is deposited between the coil and the top pole layer. The top Permalloy pole is electrodeposited in the same way as the bottom pole layer, on thin sputter-deposited Permalloy underlayer (50–100 nm). The top and bottom pole layers are in contact. Finally, Cu interconnect pads, about 25-μm thick, are electrodeposited. The entire structure, poles and coil, is protected by an overcoat, usually sputtered Al_2O_3. The dimensions of an actual fabricated write head are about 100×100 μm.

The function of the write (record) head is to generate a high magnetic field across the gap (H_g) in order to write on the high coercivity storage media. The magnetic field across the gap, H_g, increases with applied current through the Cu coil according to the equation

$$H_g = \eta \frac{N_w I}{g} \tag{91}$$

where η is the head efficiency, N_w is the number of turns in Cu coil on the write head, I is the applied current, and g is the gap length. The efficiency η increases with an increase in the relative permeability μ_r of the pole material. In the recent designs, N_w is increased to 54 turns, and typical write heads have a multilayer Cu coil structure (e.g. four layers). The gap magnetic field H_g is limited by the saturation magnetization M_s of the pole material

$$H_{g,max} \approx 0.8 \times M_s \tag{92}$$

Thus, advanced materials for write head should have high saturation magnetization. An increase in performance parameters such as the increase in bit density, that is, decrease in bit size and high data rate (high switching rate), has created a demand for new material for storage media and the magnetic material in the write head poles. Now computers have bit densities of more than 10 Gbits/inch^2 and bit sizes of about 3 μm by 0.18 μm. A bit density of 100 Gbits/inch^2 is under development. The size of the smallest bit length a is a function of the coercivity of the media H_c, the magnetization of the media M_r,

and the thickness d of the media

$$a = (d/2\pi)\frac{M_r}{H_c} \qquad (93)$$

Thus, one method of obtaining small bit sizes is to increase the coercivity (H_c) and to decrease the thickness of the medium (d). An increase in media coercivity (beyond 2000 Oe) requires an increase in the saturation magnetization M_s of the write head pole material in order to get higher coercivity $H_{g,max}$ (Eq. 92) beyond that of $Ni_{80}Fe_{20}$ Permalloy. A higher bit density results in a higher data rate. Permalloy, $Ni_{80}Fe_{20}$, has an M_s of 1.0 T. M_r cannot be reduced to decrease a in the above equation because a high M_r is required for the read process.

Pole materials in new write heads should have the following properties [100, 121–124]:

1. Saturation induction (saturation flux density): $B_s \geq 2.0$ T (B_s well above that of $Ni_{80}Fe_{20}$ Permalloy, 1 T)
2. Coercivity: $H_c < 1$ Oe
3. Magnetocrystalline anisotropy field: H_k in the range of 6 to 12 Oe
4. Permeability: $\mu_r > 1000$
5. Electrical resistivity of the order of 100 Ωcm
6. Saturation magnetostriction constant: $\lambda_s = \pm 2 \times 10^{-7}$.

All these requirements make the research for new materials difficult and challenging, but very attractive and interesting.

Three main directions in the research on new magnetic pole materials are (1) to change the composition of Permalloy; to develop a new Ni_xFe_y binary alloy; (2) to replace Ni with Co in Permalloy, thus forming a new binary alloy, Co–Fe; (3) to add a third element to binary alloys, for example, Co, Cu, or Cr, thus forming ternary alloys. The addition of Co to Permalloy, for example, results in Ni–Fe–Co alloys [124, 125], addition of Cu to Co–Fe results in Co–Fe–Cu, and addition of Cr to Co–Fe results in Co–Fe–Cr alloy. The most studied ternary alloys are Ni–Fe–Co alloys [124, 125].

It was found in approach (1) above that an increase in the Fe content of electroplated Permalloy, from 20 to about 55% increases the saturation magnetization from 1.0 T for $Ni_{80}Fe_{20}$ to 1.6 T for $Ni_{45}Fe_{55}$. $Co_{65}Ni_{12}Fe_{23}$, from approach (2), has an M_s of 2.1 T [124, 125].

Very intensive research is going on now to find material that satisfies all the above-mentioned (1–6) requirements. So far, no material satisfies all requirements.

3.8.3 Read Heads

Thin-film inductive heads may be used for reading processes also. When used in the reading process, read heads intercept the magnetic field from the recorded magnetization pattern in magnetic recording medium and convert it into electrical signals, which are then detected and decoded. A significant improvement in the reading process was achieved by the introduction of a separate read head using *magnetoresistive* (*MR*) *effect* to sense the magnetic field (data) in the magnetic storage medium.

The MR effect is a change of electrical resistivity of a metal or alloy when it is in a magnetic field. This change in resistivity due to the magnetic field can be used to sense magnetic field (data) in the magnetic medium. An MR read head is schematically shown in Fig. 43. It can be seen that the MR sensor is placed between two shields in order to reduce the pulse width of the MR sensor. The most

commonly used MR sensor material is ion beam–sputtered Permalloy ($Ni_{81}Fe_{19}$). The most commonly used material for MR head shields is the electrodeposited Permalloy. When the sensor of the read head intercepts the magnetic field from the magnetization pattern in magnetic recording medium, the resistance R of the sensor, through which a constant current I passes, changes and this results in a change in the voltage V

$$\Delta V = I\Delta R \tag{94}$$

For Permalloy and Co–Fe alloys, the decrease in resistivity is about 2–6% of the resistivity in the absence of the magnetic field, at room temperature [126].

Permalloy shows an MR effect and also an anisotropy of resistivity effect. In anisotropy of resistivity, the electrical resistivity depends on the mutual orientation of the magnetic field and the current through the sample. Both these effects are interpreted on the basis of the *two-band (two-current) model* [127–129]. According to this model, two groups of electrons carry current independently in parallel. These two groups are spin ↑ (spin up) and spin ↓ (spin down). There is different scattering of spin ↑ and spin ↓ electrons, resulting in different electrical resistivity ρ, $\rho\uparrow$, and $\rho\downarrow$. Thus, the conduction takes place by two currents (spin ↑ and spin ↓). In resistivity anisotropy, the difference in resistivity for currents perpendicular to and parallel to the magnetization direction is due to the spin-orbit-induced resistivity transfer from the spin-down electron current to the spin-up electron current [130].

Enhanced magnetoresistive effect, called *giant magnetoresistance* (GMR) *effect*, was observed in magnetic layered structures consisting of *magnetic/nonmagnetic metal multilayers* [131, 132]. First, GMR multilayers were produced by vacuum deposition technique. In the absence of a magnetic field ($H = 0$), magnetic metal layers are arranged (coupled) antiferromagnetically (AF) (Fig. 44). During the magnetization of these multilayers, the applied magnetic field aligns the magnetic moments of successive layers. In a number of multilayer systems, an antiferromagnetic (AF) arrangement is changed into a ferromagnetic (F) one (Fig. 44). The electrical resistance decreases when the film magnetizations rotate from the AF to the F alignment. Thus, the resistivity of these systems is strongly dependent on the magnetic state of the magnetic layers in the multilayer structures. The theory of this change in resistivity is based on an extended theory of electrical resistivity of metals. According to the modern quantum electronic theory, the electrical resistivity of metals results from the scattering of electrons by the lattice [89, 90, 133]. The extended theory takes into account the spin of electrons and the spin-dependent scattering in the film and at the interfaces between films [134, 135]. An image of the specimen magnetization may be obtained by scanning electron microscopy with polarization analysis (SEMPA).

The GMR multilayers require a large saturation field (>10 000 Oe) to overcome the antiferromagnetic coupling of magnetic layers in order to display the large GMR ratios. These multilayers are potentially useful for application in magnetic read heads.

Parkin et al. [136] reported saturation magnetoresistance values ($\Delta R/R$) of more than 65% at room temperature (295 K) in magnetron-sputtered [137] antiferromagnetic Co/Cu multilayers. The magnitude of magnetoresistance is usually expressed by $\Delta R(H)/R_0$, or simply $\Delta R/R$, where $R(H)$ is the resistance in field H, $R_0 = R(H = 0)$ is the sample resistance

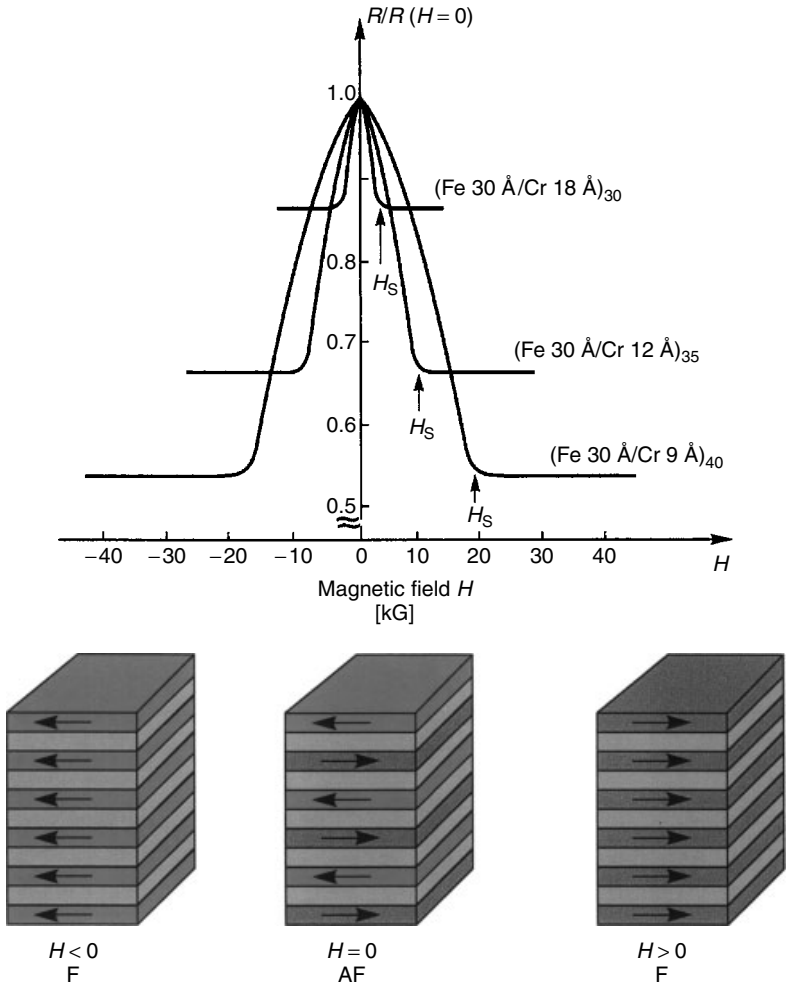


Fig. 44 Giant magnetoresistance versus field of three $(Fe/Cr)_n$ multilayers at 4.2 K. The current and the applied field are along the same [108] axis in the plane of the layers (from Ref. 131 with permission from American Physical Society).

in zero magnetic field, and $\Delta R(H) = R(H) - R_0$. Figure 45 shows $\Delta R/R$ (%) as a function of the magnetic field curves for four identical $[Co(10\ Å)/(Cu(9\ Å))]_{16}$ multilayers, consisting of 16 bilayers of Co/Cu, grown on 50 Å Cu or Fe buffer underlayers with 50 Å Fe or Cu capping (over) layers. It can be seen that the shape of the resistance curves is distinctly different, depending on the type of over- and underlayers. This effect of over- and underlayers is explained in terms of current shunting [136]. In a zero field, adjacent Co layers are aligned (coupled) antiferromagnetically with respect to one another [138].

Bird and Schlesinger [139] reported peak GMR values of 55% in electrodeposited $[Co/Cu]_n$ multilayers with $n = 600$

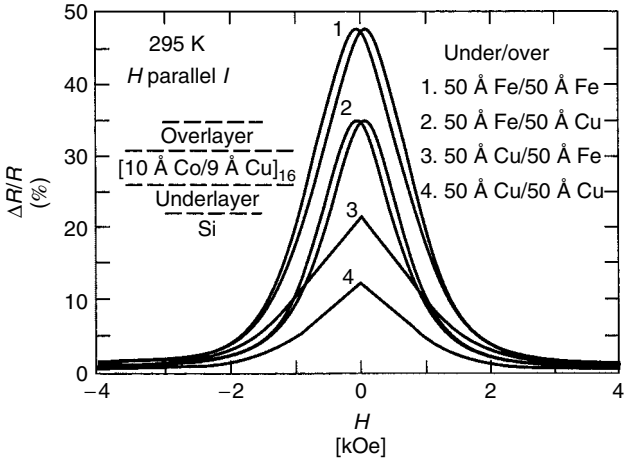


Fig. 45 Room-temperature magnetoresistance versus field curves for four different under-/overlayers (from Ref. 140 with permission from American Institute of Physics).

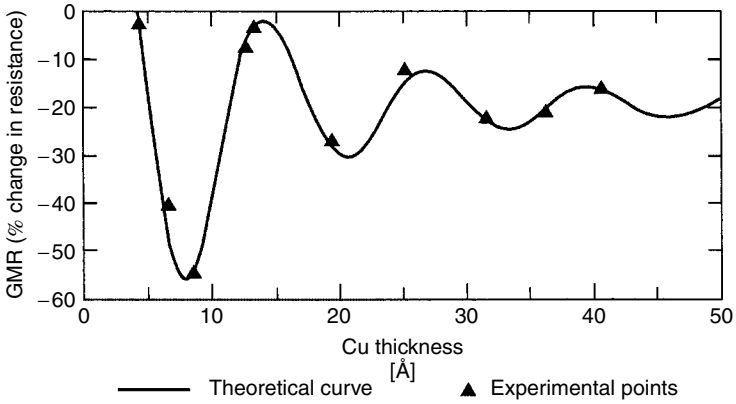


Fig. 46 Saturation magnetoresistance as a function of Cu thickness for electrodeposited $[Co/Cu]_{600}$ multilayers (from Ref. 139 with permission from the Electrochemical Society).

(Fig. 46). These multilayers were deposited potentiostatically from a single solution containing both Co and Cu ions with a Co:Cu ratio of 100:1. The potential of the cathode was pulsed between −1.40 V (versus SCE) for the cobalt deposition and −0.65 V (versus SCE) for the copper deposition with an ≈3 s cell potential interrupt between the Co-to-Cu transition. Nesbet published a critical review of theoretical studies of spin-dependent conductivity in GMR [140].

One problem of the application of GMR multilayer structures is that saturation magnetic fields are generally greater than 1000 Oe. Dieny et al. [141] eliminated this problem in 1991 by discovering the *spin valve effect*. They reported the GMR

effect in low magnetic fields, *low-field GMR*, for sandwiches consisting of two uncoupled ferromagnetic layers separated by ultrathin nonmagnetic metal. The original Permalloy spin valve structure is Si/150 Å NiFe/26 Å Cu/150 Å NiFe/100 Å FeMn/20 Å Ag. It has a GMR ratio ($\Delta R/R$) between 2 and 6%. Insertion of Co layers at each NiFe/Cu interface enhances the GMR ratio [134, 142]. After the publication of 1991, many variations of the spin valve structure were reported; for example, Anthony et al. reported a spin valve structure displaying a GMR ratio of 12% [143].

The GMR head technology was brought into production in 1998 (10 years after the discovery of the GMR effect).

There are no reports on spin valve structures built by electrodeposition. One possible difficulty of the use of electrodeposition is the ultrathin thickness of metal layers in these structures.

3.8.4 Magnetic Recording Media

There are two types of magnetic recording media: particulate media and thin-film media. *Particulate media* are composed of magnetic particles embedded in an organic polymer matrix. The most commonly used particles are γ-Fe_2O_3 (gamma ferric oxide), cobalt-modified gamma ferric oxide, γ-Fe_2O_3 + Co, CrO_2 (chromium dioxide), metal particles (Fe−Ni, Fe−Co), and barium ferrite ($BaFe_{12-2x}Co_xTi_xO_{19}$). The metal particles and barium ferrite are preferred materials for high-density applications. Particulate magnetic media are prepared by chemical precipitation techniques. The coercivity of particulate media is in the range 250 to 2300 Oe [144].

In the early 1960s, particulate media dominated the rigid disk as well as the tape industry. An increase in performance parameters such as the increase in bit density and high data rate (high switching rate) has created a demand for new material for storage media. In the late 1980s, it was realized that it is easier to get higher performance for thin-film disks than for particular disks. Thus, particulate media for rigid disks were replaced by superior *thin-film media*. Particulate media are now used in floppy drives and magnetic tapes, and almost all advanced disk drives use thin films.

Thin-film recording media evolved in three major deposition stages: electrochemical, physical–electrochemical, and physical deposition [145, 146].

1. *Electrochemical deposition*: All the early thin-film recording media were electroless or electrodeposited cobalt or cobalt alloys. Electroless Co(P) deposits contain 2–4% P. The coercivity (H_c) of these deposits is a function of the deposit thickness d according to one of the two functions, depending on deposition conditions [147–149]

$$H_c = d^{-1/3} \text{ or } H_c = d^{-2/5} \qquad (95)$$

Figure 47 shows H_c as a function of the deposit thickness for Co(P) deposited from chloride and sulfate electroless Co solutions. Fisher and Chilton [147] reported H_c values of 470 and 360 at thicknesses of 4000 and 10 000 Å, respectively, for crystalline isotropic Co(P) deposits. The demand for higher recording density, smaller a (Eq. 93), resulted in the research on thinner Co(P) deposits of higher coercivity. Maradzadeh [148] reported a coercivity of 80–900 Oe for electroless deposits of about 800 Å thick. Judge et al. reported a coercivity of 1100 Oe for thinner (600 Å) Co(P) deposits (Fig. 47).

Further increase in coercivity was achieved by the addition of Ni to Co(P)

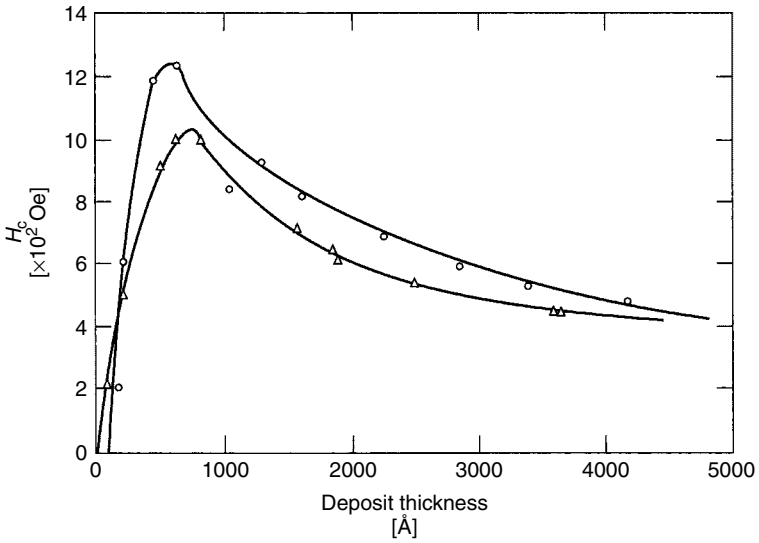


Fig. 47 Coercivity versus deposit thickness for electroless Co(P) deposits; electroless Co solutions: Δ, chloride; o, sulfate (from Ref. 149 with permission from American Institute of Physics).

and, thus, depositing electroless Co−Ni(P) alloy. Judge et al. [149] reported a peak coercivity of 1300 Oe for electroless Co−Ni(P) with 10–30% Ni and thickness of 450–680 Å.

Electrodeposited Co−Ni−(P) alloys show coercivity of 1020 Oe at the thickness of 3200 Å [150].

Figure 48 shows the general structure of electroless deposited thin-film recording medium (rigid disk from 1980s) consisting of a substrate, an underlayer, magnetic layer, and an overcoat. The substrate is an Al−Mg alloy and the underlayer is electroless Ni(P). Ni(P) should be deposited in such a way that it is nonmagnetic (see

Disk cross section	Thickness	Process
Lubricant	10 Å	Dipping
Protective layer (C)	150 Å	Sputtering
Magnetic layer (Co alloy)	~500 Å	Electroless deposition
Underlayer (Ni–P)	~10 μm	Electroless deposition
Substrate (Al–Mg alloy)	~0.6 mm	

Fig. 48 Schematic cross section of a general electroless deposited thin-film disk magnetic recording medium.

Sect. 3.6.8). Most magnetic thin-film layers are electroless Co(P) and Co−Ni(P) alloys. Overcoat is a bilayer consisting of SiO_2 and a lubricant.

2. *Physical–electrochemical deposition*: Figure 49, shows a schematic cross section of the currently dominant magnetic disk structure. The substrate is the Al−Mg or Al−Si−Cr alloy. The underlayer is a bilayer consisting of sputtered Cr and electroless Ni(P). The magnetic layer is a sputtered Co alloy. Thus, in this new

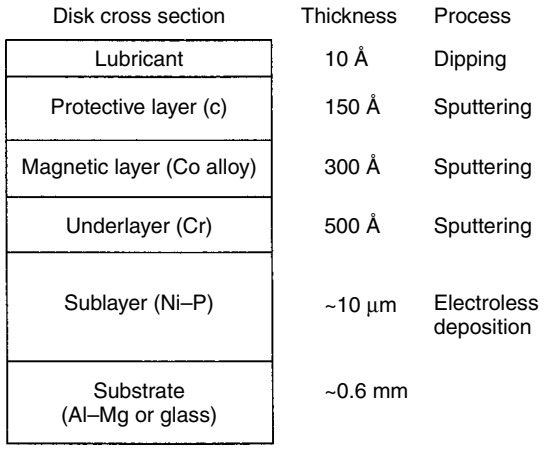


Fig. 49 Schematic cross section of a dominant thin-film magnetic disk structure.

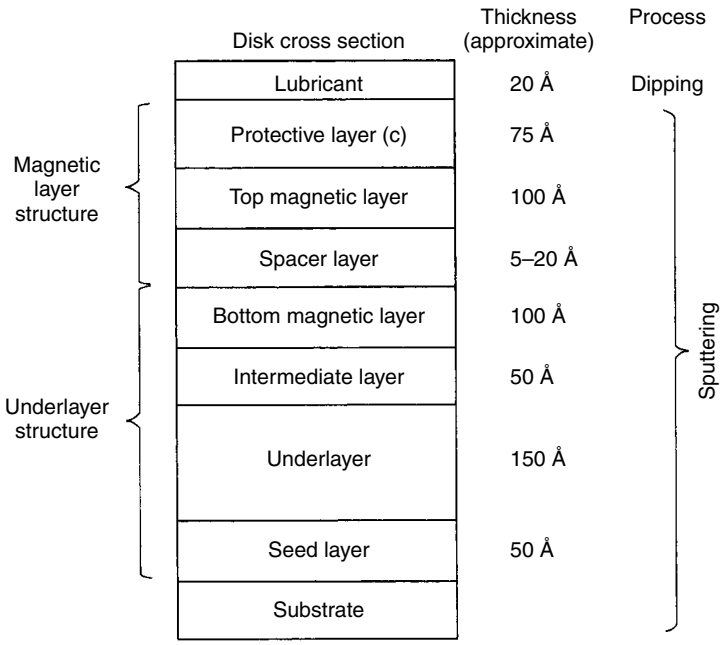


Fig. 50 Schematic cross section of a thin-film disk with complex, multilayer structures of magnetic layer and underlayer (from Ref. 157 with permission from American Institute of Physics).

technology (since the late 1980s), only one layer is electrolessly deposited. Today, electrochemically plated magnetic thin films are no longer used in manufacturing. A chromium underlayer is introduced to control the microstructure and magnetic properties of the Co alloy magnetic recording layer [151–153]. The dominant overcoat is sputtered hard carbon (DLC, diamond-like carbon).

Concerns about the corrosion of thin films resulted in the evaluation of Co–Pt alloys. Sputtered Co–Pt magnetic thin films containing 25% Pt and 500 Å thick have a coercivity of 1700 Oe [154]. A coercivity higher than 3000 Oe was obtained with Co–Cr–Pt–B sputtered media [155] on 3000 Å thick Cr underlayer (no intermediate layer is required). Boron, the fourth element, is added at an amount of about 3 to 7 at%. A coercivity of 4200 Oe was obtained for sputtered Co–Cr–Pt alloys for a thickness range between 100 and 350 Å. These films require a Cr underlayer and a special intermediate layer composed of Co or Co–Cr–Ta [156].

3. *Physical deposition*: Recent disks with glass substrate do not have the Ni(P) underlayer [156]. Thus, the last electroless layer, Ni(P), is eliminated in some products. The underlayers considered currently have complex structures composed of up to three layers. Thin-film magnetic layers are complex and composed of three layers [157], Fig. 50.

Currently it appears that further research on electrochemical deposition of magnetic thin films is of academic interest only. This research can be very useful in developing new structures and new ideas that can be later adopted by physical techniques of deposition. Recent developments in this area include work on longitudinal and perpendicular magnetic recording media [158, 159].

Thompson and Best discuss the future of magnetic data storage technology [120].

References

1. M. Paunovic, M. Schlesinger, *Fundamentals of Electrochemical Deposition*, 2nd ed., Wiley, New York, 2006.
2. C. Kittel, *Introduction to Solid State Physics*, Wiley, New York, 1996.
3. H. Eyring, J. E. Walter, G. E. Kimball, *Quantum Chemistry*, Wiley, New York, 1944.
4. S. Glasstone, K. J. Laidler, H. Eyring, *The Theory of Rate Processes*, McGraw-Hill, New York, 1941.
5. T. Erdey-Gruz, M. Volmer, *Z. Phys. Chem.* **1930**, *A150*, 203–213.
6. W. Lorenz, *Z. Phys. Chem.* **1953**, *202B*, 275–291.
7. J. O. M. Bockris, in *Transactions of the Symposium on Electrode Processes* (Ed.: E. Yeager), Wiley, New York, 1961.
8. W. Lorenz, *Z. Elektrochem.* **1953**, *57*, 382–385.
9. W. Lorenz, *Naturwissenschaften* **1953**, *40*, 576–577.
10. W. J. Lorenz, *Z. Naturforsch.* **1954**, *9A*, 716–724.
11. W. Mehl, J. O. M. Bockris, *J. Chem. Phys.* **1957**, *27*, 817–824.
12. B. E. Conway, J. O. M. Bockris, *Proc. R. Soc.* **1958**, *A248*, 394–403.
13. B. E. Conway, J. O. M. Bockris, *Electrochim. Acta* **1961**, *3*, 340–354.
14. G. H. Gilmer, P. Bennema, *J. Appl. Phys.* **1972**, *43*, 1347–1360.
15. (a) A. J. Bard, L. R. Faulkner, *Electrochemical Methods*, Wiley, New York, 1980. (b) D. D. Macdonald, *Transient Techniques in Electrochemistry*, Plenum Publishing, New York, 1977.
16. E. Budevski, G. Staikov, W. J. Lorenz, *Electrochemical Phase Formation and Growth*, VCH, New York, 1996.
17. M. A. Schneeweiss, D. M. Kolb, *Phys. Status Solidi A* **1999**, *173*, 51–71.
18. D. M. Kolb, R. Ullmann, J. C. Ziegler, *Electrochim. Acta* **1998**, *43*, 2751–2760.
19. M. Giesen, R. Randler, S. Baier et al., *Electrochim. Acta* **1999**, *45*, 527–536.
20. Q. Wu, D. Barkey, *J. Electrochem. Soc.* **1997**, *144*, L261–L262.

21. J. C. Vickerman, *Surface Analysis – the Principal Techniques*, Wiley, New York, 1998.
22. T. L. Hill, *Statistical Thermodynamics*, Addison-Wesley, Reading, 1962.
23. R. Ullmann, T. Will, D. M. Kolb, *Chem. Phys. Lett.* **1993**, *209*, 238–242.
24. G. E. Engelmann, J. C. Ziegler, D. M. Kolb, *Surf. Sci.* **1998**, *401*, L420–L424.
25. R. D. Armstrong, J. A. Harrison, *J. Electrochem. Soc.* **1969**, *116*, 328–334.
26. B. Scharifker, G. Hills, *Electrochim. Acta* **1983**, *28*, 879–889.
27. R. M. Rynders, R. Alkire, *J. Electrochem. Soc.* **1994**, *141*, 1166–1173.
28. V. Bostanov, G. Staikov, D. K. Roe, *J. Electrochem. Soc.* **1975**, *122*, 1301–1305.
29. F. C. Frank, in *Growth and Perfection of Crystals* (Eds.: R. H. Doremus, B. W. Roberts, D. Turnbull), Wiley, New York, 1958.
30. G. Wranglen, *Electrochim. Acta* **1960**, *2*, 130–144.
31. A. Damjanovic, M. Paunovic, J. O. M. Bockris, *J. Electroanal. Chem.* **1965**, *9*, 93–111.
32. U. Bertocci, C. Bertocci, *J. Electrochem. Soc.* **1971**, *118*, 1287–1298.
33. E. W. Dickinson, M. H. Jacobs, D. W. Pashley, *Philos. Mag.* **1965**, *11*, 575–590.
34. R. Sard, R. Weil, *Electrochim. Acta* **1970**, *15*, 1977–1989.
35. S. Nakahara, R. Weil, *J. Electrochem. Soc.* **1973**, *120*, 1462–1469.
36. S. Nakahara, *Thin Solid Films* **1977**, *45*, 421–432.
37. R. Weil, J. B. C. Wu, *Plating* **1973**, *60*, 622–626.
38. J. O. Dukovic, C. Tobias, *J. Electrochem. Soc.* **1990**, *137*, 3748–3755.
39. C. Madore, D. Landolt, *J. Electrochem. Soc.* **1996**, *143*, 3939–3943.
40. O. Kardos, D. G. Foulke, in *Advances in Electrochemistry and Electrochemical Engineering* (Ed.: C. W. Tobias), Interscience Publishers, New York, 1962, Vol. 2.
41. L. Onicin, L. Muresan, *J. Appl. Electrochem.* **1991**, *21*, 565–573.
42. G. T. Rogers, K. J. Taylor, *Electrochim. Acta* **1963**, *8*, 887–904; **1966**, *11*, 1685–1696; **1968**, *13*, 109–117.
43. M. Paunovic, *Plating* **1968**, *55*, 1161–1167.
44. M. Saito, *J. Met. Finish. Soc. Jpn.* **1966**, *17*, 14–34.
45. C. Wagner, W. Traud, *Z. Electrochem.* **1938**, *44*, 391–401.
46. F. M. Donahue, *Oberflache-Surf.* **1972**, *13*, 301–308.
47. A. Molenaar, M. F. Holdrinet, L. K. H. van Beek, *Plating* **1974**, *61*, 238–242.
48. S. M. El-Raghy, A. A. Abo-Salama, *J. Electrochem. Soc.* **1979**, *126*, 171–176.
49. M. Paunovic, *Plat. Surf. Finish.* **1983**, *70*, 62–68.
50. Y. Okinaka, *J. Electrochem. Soc.* **1973**, *120*, 739–744.
51. M. Paunovic, R. Arndt, *J. Electrochem. Soc.* **1983**, *130*, 794–799.
52. D. Vitkavage, M. Paunovic, *Plat. Surf. Finish.* **1983**, *70*(4), 48–50.
53. J. E. A. M. Van den Meerakker, *J. Appl. Electrochem.* **1981**, *11*, 395–400.
54. R. P. Buck, L. R. Griffith, *J. Electrochem. Soc.* **1962**, *109*, 1005–1013.
55. M. Paunovic, *J. Electrochem. Soc.* **1978**, *125*, 173–174.
56. M. Charbonnier, M. Alami, M. Romand, *J. Electrochem. Soc.* **1996**, *143*, 472–480.
57. M. Paunovic, *J. Electrochem. Soc.* **1980**, *127*, 441C.
58. J.-Y. Zhang, H. Esrom, I. W. Boyd, *Appl. Surf. Sci.* **1996**, *96–98*, 399.
59. C. H. Ting, M. Paunovic, *J. Electrochem. Soc.* **1989**, *136*, 456–462.
60. G. J. Norga, M. Platero, K. A. Black et al., *J. Electrochem. Soc.* **1997**, *144*, 2801–2810.
61. J.-Y. Zhang, I. W. Boyd, H. Esrom, *Appl. Surface Sci.* **1997**, *109/110*, 253–260.
62. N. Massee, D. L. Piron, *J. Electrochem. Soc.* **1993**, *140*, 2818–2824.
63. L. E. Murr, V. Annamalai, *Metall. Mater. Trans.* **1978**, *9B*, 515–520.
64. K. M. Gorbunova, Yu. M. Polukarov, in *Advances in Electrochemistry and Electrochemical Engineering* (Ed.: C. W. Tobias), Wiley, New York, 1967, Vol. 5.
65. S. Glasstone, T. E. Symes, *Trans. Faraday Soc.* **1927**, *23*, 213–226.
66. J. O. M. Bockris, D. Drazic, A. R. Despic, *Electrochim. Acta* **1961**, *4*, 325–361.
67. J. Matulis, R. Slizys, *Electrochim. Acta* **1964**, *9*, 1177–1182.
68. H. Dahms, I. M. Croll, *J. Electrochem. Soc.* **1965**, *112*, 771–775.
69. L. Harris, *J. Electrochem. Soc.* **1973**, *120*, 1034–1039.
70. S. Hessami, C. W. Tobias, *J. Electrochem. Soc.* **1989**, *136*, 3611–3616.
71. W. C. Grande, J. B. Talbot, *J. Electrochem. Soc.* **1993**, *140*, 669–675.

72. K. Y. Sasaki, J. B. Talbot, *J. Electrochem. Soc.* **1995**, *142*, 775–782; **2000**, *146*, 189–197.
73. D. Golodnitsky, N. V. Gudin, G. A. Volyanuk, *J. Electrochem. Soc.* **2000**, *147*, 4156–4163.
74. I. Ohno, O. Wakabayashi, S. Haruyama, *J. Electrochem. Soc.* **1985**, *132*, 2323–2330.
75. G. E. Georgiou, F. A. Baiocchi, H. S. Luftman et al., *J. Electrochem. Soc.* **1991**, *138*, 2061–2069.
76. Y. Shacham-Diamand, S. Lopatin, *Microelectron. Eng.* **1997**, *37/38*, 77–78.
77. H. Sawai, T. Kanamori, I. Koiwa et al., *J. Electrochem. Soc.* **1990**, *137*, 3653–3660.
78. T. Osaka, T. Homma, T. Yokoshima, in *Modern Electroplating*, 4th ed. (Eds.: M. Schlesinger, M. Paunovic), Wiley, New York, 2000, Chapter 19.
79. I. Ohno, in *Modern Electroplating*, 4th ed. (Eds.: M. Schlesinger, M. Paunovic), Wiley, New York, 2000, Chapter 22.
80. H. Seiter, H. Fischer, L. Albert, *Electrochim. Acta* **1960**, *2*, 97–120.
81. S. C. Barnes, G. G. Storey, H. J. Pick, *Electrochim. Acta* **1960**, *2*, 195–206.
82. D. J. Srolovitz, A. Mazor, G. G. Bukiet, *J. Vac. Sci. Technol.* **1988**, *A6*, 2371–2380.
83. G. L. M. K. S. Kahandra, X. Zou, R. Farrell et al., *Phys. Rev. Lett.* **1992**, *56*, 3741.
84. M. Amblart, G. Froment, N. Maurin et al., *Electrochim. Acta* **1983**, *28*, 909–915.
85 A.-L. Barabasi, H. E. Stanly, *Fractal Concepts in Surface Growth*, Cambridge University Press, Cambridge, 1995.
86. P. B. Price, D. A. Vermilyea, M. B. Webb, *Acta Metall. Mater.* **1958**, *6*, 524–531.
87. Y. Zhang, J. A. Abys, in *Modern Electroplating*, 4th ed. (Eds.: M. Schlesinger, M. Paunovic), Wiley, New York, 2000.
88. Z. Fisk, G. W. Webb, in *Treatise on Material Science and Technology* (Ed.: F. W. Fradin), Academic Press, New York, 1981, p. 297, Vol. 21.
89. P. L. Rossiter, *The Electrical Resistivity of Metals and Alloys*, Cambridge University Press, Cambridge, 1987.
90. M. Paunovic, L. A. Clevenger, J. Gupta et al., *J. Electrochem. Soc.* **1993**, *140*, 2690–2694.
91. J. M. E. Harper, E. G. Colgan, C.-K. Hu et al., *MRS Bull.* **1994**, *19*(8), 23–29.
92. D. S. Gardner, J. Omuki, K. Kudoo et al., *Thin Solid Films* **1995**, *262*, 104–119.
93. A. J. Learn, *J. Electrochem. Soc.* **1976**, *123*, 894–906.
94. C.-K. Hu, K. P. Rodbell, T. D. Sullivan et al., *IBM J. Res. Dev.* **1995**, *39*, 465.
95. M. Ohring, *Reliability and Failure of Electronic Materials and Devices*, Academic Press, New York, 1998, Chapter 5.
96. Y. Taur, T. H. Ning, *Fundamentals of Modern VLSI Devices*, Cambridge University Press, Cambridge, UK, 1998, Chapters 4 and 5.
97. S. Vaidya, A. K. Sinha, *Thin Solid Films* **1981**, *75*, 253–259.
98. S. P. Murarka, R. J. Gutmann, A. E. Kaloyeros et al., *Thin Solid Films* **1993**, *236*, 257.
99. (a) P. V. Zant, *Microchip Fabrication*, 4th ed., McGraw-Hill, New York, 2000, Chapter 10. (b) S. Wolf, R. N. Tauber, *Silicon Processing for VLSI Era, Volume 1-Process Technology*, 2nd ed., Lattice Press, Sunset Beach, 2000. (c) S. Wolf, *Silicon Processing for VLSI Era, Volume 2-Process Integration*, Lattice Press, Sunset Beach, 1990.
100. L. T. Romankiw, *Electrochim. Acta* **1997**, *42*, 2985–3005.
101. V. M. Dubin, Y. Shacham-Diamand, B. Zhao et al., *J. Electrochem. Soc.* **1997**, *144*, 898–908.
102. C. H. Ting, M. Paunovic, P. L. Pai et al., *J. Electrochem. Soc.* **1989**, *136*, 462–466.
103. Y. Shacham-Diamand, V. M. Dubin, *Microelectron. Eng.* **1997**, *33*, 47–58.
104. P. C. Andricacos, C. Uzoh, D. O. Dukovic et al., *IBM J. Res. Dev.* **1998**, *42*, 567–573.
105. E. K. Yung, L. T. Romankiw, R. C. Alkire, *J. Electrochem. Soc.* **1989**, *136*, 206.
106. S. Mehdizadeh, J. O. Dukovic, P. C. Andricacos et al., *J. Electrochem. Soc.* **1992**, *139*, 78–79.
107. (a) J. O. Dukovic, *IBM J. Res. Dev.* **1990**, *34*, 693–705; **1993**, *37*, 125. (b) T. J. Pricer, M. J. Kushner, R. C. Alkire, *J. Electrochem. Soc.* **2002**, *149*, C406.
108. S. Mehdizadeh, J. Dukovic, P. C. Andricacos et al., *J. Electrochem. Soc.* **1990**, *137*, 110–117.
109. R. Stolt, A. Charai, F. M. D'eurle et al., *J. Vac. Sci. Technol.* **1991**, *A9*, 1501–1509.
110. L. A. Clevenger, N. A. Bojarczuk, K. Holloway et al., *J. Appl. Phys.* **1993**, *73*, 300–308.
111. M. T. Wang, Y. C. Lin, M. C. Chen, *J. Electrochem. Soc.* **1998**, *145*, 2538–2542.

112. T. Kouno, H. Niwa, M. Yamada, *J. Electrochem. Soc.* **1998**, *145*, 2164–2170.
113. S.-Q. Wang, *MRS Bull.* **1994**, *19*, 30–41.
114. M. Paunovic, P. J. Bailey, R. G. Schad et al., *J. Electrochem. Soc.* **1994**, *141*, 1843–1850.
115. E. G. Colgan, *Thin Solid Films* **1995**, *262*, 120–125.
116. E. M. Zielinski, *J. Electron. Mater.* **1995**, *24*, 1485–1490.
117. D. P. Tracy, D. B. Knorr, K. P. Rodbell, *J. Appl. Phys.* **1994**, *76*, 2671–2680.
118. D. P. Tracy, D. B. Knorr, *J. Electron. Mater.* **1993**, *22*, 611–618.
119. P. A. Flinn, *J. Mater. Res.* **1991**, *16*, 1438–1452.
120. D. A. Thompson, J. S. Best, *IBM J. Res. Dev.* **2000**, *44*, 311–319.
121. S. X. Wang, A. M. Taratorin, *Magnetic Information Storage Technology*, Academic Press, New York, 1999.
122. R. L. Comstock, *Introduction to Magnetism and Magnetic Recording*, Wiley, New York, 1999.
123. R. E. Jones Jr., C. D. Mee, C. Tsang, in *Magnetic Recording Technology* (Eds.: C. D. Mee, E. D. Daniel), McGraw-Hill, New York, 1996, Chapter 6.
124. T. Osaka, *Electrochim. Acta* **1999**, *44*, 3885–3890.
125. T. Osaka, *Electrochim. Acta* **2000**, *45*, 3311–3321.
126. H. C. Van Elst, *Physica* **1959**, *25*, 708–720.
127. A. Fert, I. A. Campbell, *J. Phys. F. Metal. Phys.* **1976**, *6*, 848–871.
128. I. A. Campbell, A. Fert, A. R. Pomeroy, *Philos. Mag.* **1967**, *15*, 977–982.
129. T. Farrell, D. Greig, *J. Phys. C: Solid State Phys.* **1970**, *3*, 138–145.
130. I. A. Campbell, A. Fert, O. Jaoul, *J. Phys. C: Solid State Phys.* **1970**, *3*, S95–S10.
131. M. N. Baibich, J. M. Broto, A. Fert et al., *Phys. Rev. Lett.* **1988**, *61*, 2472–2475.
132. S. S. P. Parkin, *IBM J. Res. Dev.* **1998**, *42*, 3–6.
133. Z. Fisk, G. W. Webb, in *Treatise on Material Science and Technology* (Ed.: F. Y. Fradin), Academic Press, New York, 1981, p. 297, Vol. 21.
134. (a) S. S. P. Parkin, *Phys. Rev. Lett.* **1993**, *71*, 1641–1644. (b) S. S. P. Parkin, A. Modak, D. J. Smith, *Phys. Rev. B* **1993**, *47*, 9136–9139.
135. D. T. Pierce, J. Unguris, R. J. Celotta et al., in *Ultrathin Magnetic Structures II* (Eds.: B. Heinrich, J. A. C. Bland), Springer-Verlag, New York, 1994.
136. (a) S. S. P. Parkin, R. Bhadra, K. P. Roche, *Phys. Rev. Lett.* **1991**, *66*, 2152–2155. (b) S. S. Parkin, Z. G. Li, D. J. Smith, *Appl. Phys. Lett.* **1991**, *58*, 2710–2712.
137. A. E. Wendt, in *Ionized Physical Vapor Deposition* (Ed.: J. A. Hopwood), Academic Press, New York, 2000.
138. E. Fawcett, *Rev. Mod. Phys.* **1988**, *60*, 209–212.
139. K. D. Bird, M. Schlesinger, *J. Electrochem. Soc.* **1995**, *142*, L65–L66.
140. R. K. Nesbet, *IBM J. Res. Dev.* **1998**, *42*, 53–71.
141. B. Dieny, V. S. Speriosu, S. S. Parkin et al., *Phys. Rev.* **1991**, *B43*, 1297–1300.
142. S. S. P. Parkin, *Appl. Phys. Lett.* **1992**, *61*, 1358–1360.
143. T. C. Anthony, J. A. Brug, S. Zhang, *IEEE Trans. Magn.* **1994**, *30*, 3819–3821.
144. E. Koester, in *Magnetic Recording Technology* (Eds.: C. D. Mee, E. D. Daniel), McGraw-Hill, New York, 1996, Chapter 3.
145. R. A. Powell, S. M. Rossnagel, *Thin Films, PVD for Microelectronics: Sputter Deposition Applied to Semiconductor Manufacturing*, Academic Press, New York, 1999, Vol. 26.
146. J. A. Hopwood, (Eds.), *Thin Films, Ionized Physical Vapor Deposition*, Academic Press, New York, 2000, Vol. 27.
147. R. D. Fisher, W. H. Chilton, *J. Electrochem. Soc.* **1962**, *109*, 485–490.
148. Y. Moradzadeh, *J. Electrochem. Soc.* **1965**, *112*, 891–896.
149. J. S. Judge, J. R. Morrison, D. E. Speliotis, *J. Appl. Phys.* **1965**, *536*, 948–949.
150. F. E. Luborsky, *IEEE Trans. Magn.* **1970**, *MAG-6*, 502–506.
151. C. A. Ross, F. M. Ross, G. Bertero et al., *IEEE Trans. Magn.* **1998**, *34*, 282–292.
152. T. Chen, T. Yamashita, *IEEE Trans. Magn.* **1988**, *24*, 2629–2634.
153. T. Yogi, C. Tsang, T. A. Nguyen et al., *IEEE Trans. Magn.* **1990**, *26*, 2271–2276.
154. J. A. Aboaf, S. R. Herd, E. Klokholm, *IEEE Trans. Magn.* **1983**, *MAG-19*, 1514–1519.
155. N. Tani, T. Takahashi, M. Hashimoto et al., *IEEE Trans. Magn.* **1991**, *27*, 4736–4738.
156. L. L. Z. Fang, D. N. Lambeth, *Appl. Phys. Lett.* **1994**, *65*, 3137–3139.
157. K. E. Johnson, *J. Appl. Phys.* **2000**, *87*, 5365–5370.

158. T. Osaka, T. Homma, T. Yokoshima, in *Modern Electroplating* (Eds.: M. Schlesinger, M. Paunovic), Wiley, New York, 2000, Chapter 19.
159. L. Callegaro, E. Puppin, P. L. Cavallotti et al., *J. Magn. Magn. Mater.* **1996**, *155*, 190–198.
160. V. Dubin, Y. Shacham-Diamand, B. Zhao et al., *Mater. Res. Soc. Symp. Proc.* **1996**, *427*, 179–182.

Appendix

A1
Magnetization and Hysteresis Loops for Hard Magnetic Material

Consider a torus of Co–Ni(P) alloy wrapped with a coil of insulated copper wire as shown in Fig. 51. When the current I on the torus is first turned on, the magnetization M increases with increasing H along the curve **a** (Fig. 52). At higher values of H, the magnetization curve levels off, remaining nearly constant with further increase in H. This maximum magnetization of a magnetic material, when it is fully saturated, is called *saturation magnetization* M_s in the positive direction. Curve a is called the *magnetization curve*.

When, after reaching saturation magnetization, we decrease the current in the torus coil to bring H back to zero, the magnetization M decreases along curve b. It can be seen from Fig. 52 that when H reaches zero, there is still some magnetization M left. Thus, even with no magnetizing current, after the applied field is removed, there is a magnetization on the torus. The torus has become permanently magnetized. This remaining magnetization, after the applied field is removed, is called *remanence magnetization* M_r. If we now turn on a negative current in the coil, the $M-H$ curve continues along b curve until the torus is magnetically saturated in the negative direction. This maximum magnetization is called *saturation magnetization* – M_s in the negative direction.

If, after reaching saturation in the negative direction, we decrease the current in the coil to bring H back to zero again, magnetization M goes along curve c. Again, there is some M left after the applied field is removed. This magnetization is called *remanence magnetization* – M_r. With a further increase in current and H, the $M-H$ curve continues along curve c until M_s is reached. Curves b and c form a *hysteresis loop*. In general, this loop is symmetrical.

Two important points in the hysteresis loop are the intercepts of b and c curves with the H axes. These intercepts are called *coercivity*, $-H_c$ *and* $+H_c$ (or coercive force H_c). The coercivity is defined as the magnetic field needed to reduce the magnetization M to zero, starting from a saturated condition; or the magnetic field needed to switch the direction of magnetization. Magnetic materials with $H_c < 10$ Oe are called *soft magnetic materials*, while those with $H_c > 100$ Oe are called *hard magnetic materials*. The hard magnetic materials are

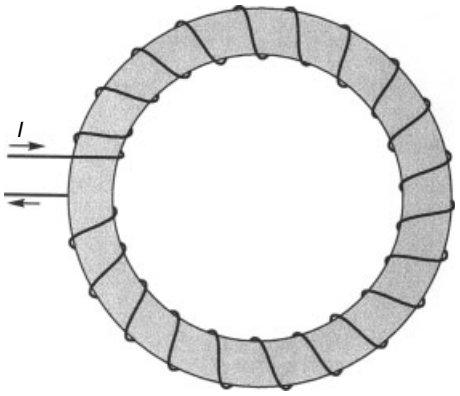


Fig. 51 A torus of Co–Ni(P) alloy wound with a coil of insulated copper wire.

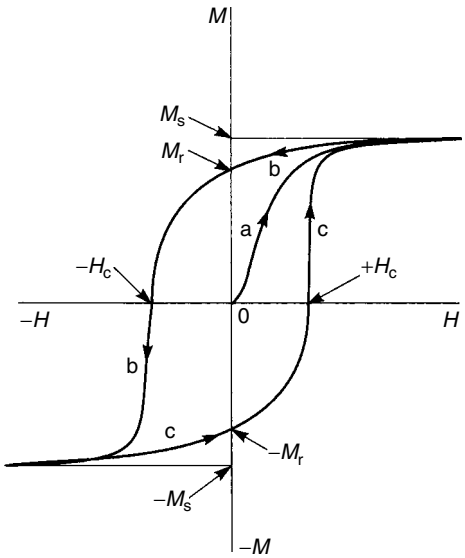


Fig. 52 Hysteresis loop of hard magnetic material.

used for magnetic recording media since their high H_c values prevent them from demagnetization.

In digital computers, the two magnetization states $+M_r$ and $-M_r$ represent elements 0 and 1 of magnetic memory. The coercive force H_c is the field required to switch from $+M_r$ to $-M_r$, or vice versa.

A2
Magnetization and Hysteresis Loops for Soft Magnetic Material

Figure 53 shows a hysteresis loop for a soft magnetic material. It is very narrow in

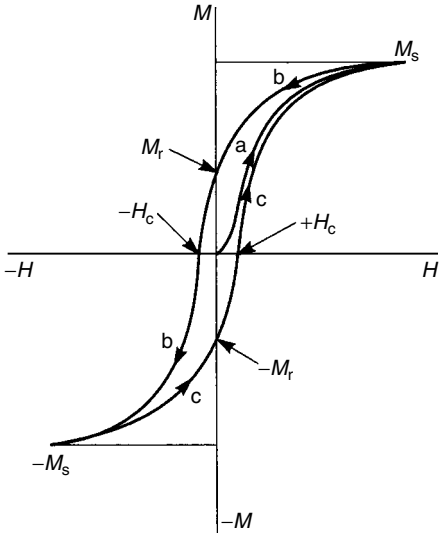


Fig. 53 Hysteresis loop of a soft magnetic material.

comparison with that of a hard material. In this case, there is a nearly perfect backtrack as the magnetic field is removed. It is very easy to magnetize a soft magnetic material such as Permalloy. Soft magnetic materials are used for thin-film heads.

A3
Magnetic Anisotropy

The magnetization, $M = f(H)$, depends on the direction of the magnetization with respect to the crystal axes in a single crystal or in a domain in a polycrystal. Figure 54 shows the magnetization curves for a single crystal of hexagonal-close-packed (hcp) Co metal when magnetized along (1) the hcp symmetry or c axis and (2) the basal plane. It can be seen that the $M = f(H)$ curve along the c axis [0001] reaches saturation at a value of H of approximately 1 kOe, while that along the basal plane does not reach saturation even at 10 kOe. The c axis is referred to as the magnetocrystalline *easy axis* and any direction in the basal plane is referred to as the *hard axis*.

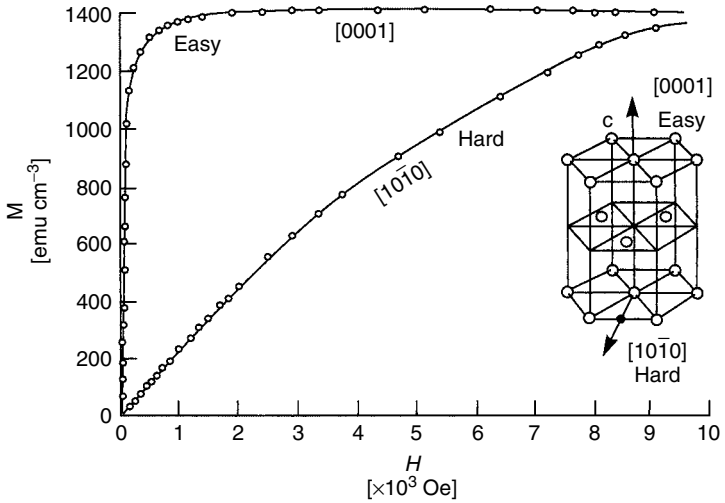


Fig. 54 Magnetization curve for easy [0001 — *c* axis] and hard [1010] axis for cobalt showing large magnetocrystalline anisotropy (from Ref. 122).

This effect is called *magnetocrystalline anisotropy*.

A4 Magnetostriction

Magnetostriction is defined as the change in dimensions of a magnetic sample during magnetization. This change is described by the magnetostriction constant λ, which is defined as the fractional change in length along the magnetization

$$\lambda = \frac{\Delta l}{l} \tag{96}$$

where l is the length of a material. When the sample is magnetized from a demagnetized state to saturation, the saturation magnetostriction constant λ_s is obtained.

4.1 Aqueous Processing of Metals

D.Sc. Jari Aromaa
Helsinki University of Technology, TKK, Finland

4.1.1 Electrowinning and Electrorefining

Electrowinning, electrorefining, and electroplating involve the exchange of electrons between a solid electrode and ions or molecules dissolved in solution. The underlying physical and chemical laws are the same for most metals and applications. The rate of the reaction involved depends on the electrode potential, the electrode area, and the rate of mass transport of the electroactive species to the electrode surface. The main differences are in the construction of the cells, their geometry, construction materials, and operating practice. The process examples presented have been selected on the basis of their historical or technological importance.

The background of electrolysis processes is the Faraday's law that connects the amount of electricity and mass of reacted material. The quantity of electricity in coulombs equals the number of electrons involved in the reaction multiplied by the Faraday's constant (F). The value of Faraday's constant is 96 485 C, which is the quantity of electricity of a mole of electrons (6.023×10^{23} electrons). The quantity of electrical current in coulombs that passes through a conductor is proportional to both the current and the length of time it flows. Thus, amperes multiplied by time gives coulombs or $q = I\,(t_2 - t_1)$ when the current is constant during the time interval between t_1 and t_2. Assuming a constant current and a 100% current efficiency, this expression can be written as Faraday's law (1)

$$\Delta m = \frac{I \cdot \Delta t}{z \cdot F} \tag{1}$$

where

- Δm is mass (g)
- I is current (A)
- Δt is time (s)
- M is atomic weight (g mol^{-1})
- z is valency of the metal (number of electrons)
- F is Faraday's constant (96 485 C mol^{-1}).

The current efficiency is the yield based on the electrical charge passed during electrolysis. If the current efficiency is less than 100%, then Eq. (1) is modified as (2)

$$\Delta m = \frac{I \cdot \Delta t}{z \cdot F} \cdot \eta \tag{2}$$

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

where

η is current efficiency (%).

A value of current efficiency below 100% usually indicates that by-products are formed. Another possibility is that reverse reaction happens, for example, dissolution of deposited metal. In metal recovery from acidic solutions, the hydrogen evolution reaction is a common side reaction that lowers the current efficiency. The current efficiency can also vary with the current density. Current efficiency is a widely used measure of tankhouse proficiency in producing metal, but it does not provide a direct measure of metal quality.

Using Faraday's law it is possible to calculate a so-called electrochemical equivalent for a metal as in Eq. (3)

$$ekv(\mathrm{g\ A^{-1}S^{-1}}) = \frac{M}{z \cdot F} \qquad (3)$$

Table 1 lists electrochemical equivalents for some metals. For example, to deposit 1 g of copper it is necessary to use $2 \times 96\,485$ As $\mathrm{mol}^{-1}/63.5$ g $\mathrm{mol}^{-1} = 3039$ As or 0.844 Ah. By using the electrochemical equivalent, it is possible to calculate the current density and time required to deposit a certain weight of metal.

Figure 1 shows the principle of electrolysis processes. The electrolysis cell consists of two electrodes (anode and cathode) and an electrolyte. In an electrolysis cell such as in electrowinning or refining, the negative electrode is the cathode and the positive electrode is the anode. The electrodes are electrically conductive materials that are in connection with the electrolyte, and the electrochemical reactions take place on the boundary between electrode and electrolyte. When the two electrodes are placed in a solution containing metal ions and an electric current is passed between them, the metal can be deposited on the cathode. The electrolyte next to the anode is the anolyte and that next to the cathode is the catholyte. The properties of anolyte and catholyte differ from those of the bulk electrolyte.

In electrorefining, impure metal is dissolved from the anode and pure metal is

Tab. 1 Electrochemical equivalents of some metals

Metal	*Molar mass [g mol^{-1}]*	*Valence*	*Electrochemical equivalent [mg A^{-1}s^{-1}]*	*Electrochemical equivalent [g A^{-1}h^{-1}]*
Ag	107.9	1	1.12	4.03
Au	197.0	1	2.04	7.34
Au	197.0	3	0.68	2.45
Cd	112.4	2	0.58	2.09
Co	58.9	2	0.31	1.12
Cu	63.5	1	0.66	2.38
Cu	63.5	2	0.33	1.19
Ni	58.7	2	0.30	1.08
Pb	207.2	2	1.07	3.85
Pb	207.2	4	0.54	1.94
Pt	195.1	2	1.01	3.64
Sn	118.7	2	0.62	2.23
Sn	118.7	4	0.31	1.12
Zn	65.4	2	0.34	1.22

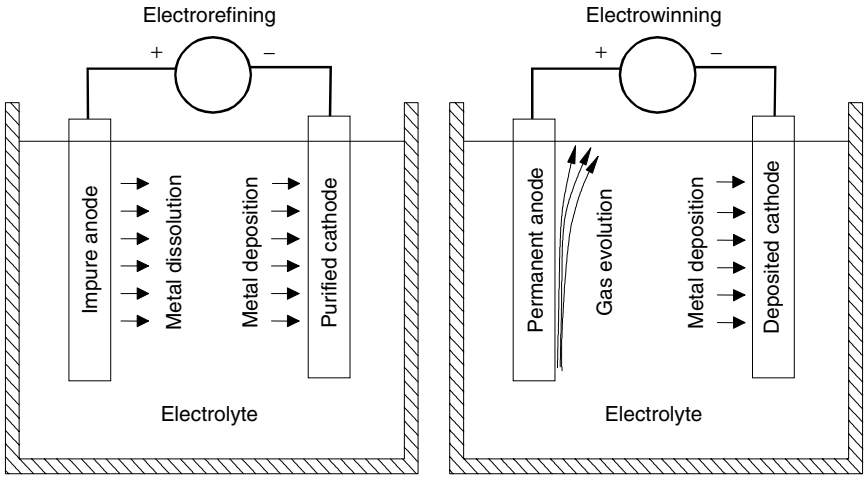


Fig. 1 The principle of electrolysis processes.

deposited on the cathode. Electrorefining is a successful method because metallic impurities that are more noble than the metal to be refined do not dissolve into the electrolyte. These impurities often include precious metals, silver, gold, and the platinum-group metals (PGMs), which are recovered by treatment of the sludge that falls to the bottom of the cell. Since impurities that do enter the electrolyte are less noble than the material to be refined, they do not deposit on the cathode. Some metals accumulate in the electrolyte and are extracted in a separate stage. Electrolyte, current density, and additives are chosen to produce a dense, compact deposit.

Electrowinning is done from purified solutions. The anodes are insoluble metal or oxide. In the electrowinning of most metals, oxygen is evolved from decomposition of water at the anode. Electrowinning of some metals is done from chloride solutions with evolution of chlorine gas as anode reaction. The metal to be recovered is deposited on the cathode as in the electrorefining process.

In electrowinning, the common anodic reactions are oxygen evolution (4) and chlorine evolution (5). Oxygen evolution by decomposition proceeds in most acid solutions, whereas chlorine evolution proceeds in chloride solutions.

$$2H_2O \longrightarrow O_2 + 4H^+ + 4e^- \quad (4)$$

$$2Cl^- \longrightarrow Cl_2 + 2e^- \quad (5)$$

In electrorefining, the anodic reaction is oxidation and dissolution of metal by the generic reaction (6). The cathodic reaction in both electrowinning and electrorefining is the reduction and deposition of metal ion by the generic reaction (7)

$$Me \longrightarrow Me^{z+} + ze^- \quad (6)$$

$$Me^{z+} + ze^- \longrightarrow Me \quad (7)$$

Table 2 lists standard electrode potentials for some metals and some other reactions common in electrorefining and electrowinning. The metals with high equilibrium potential are noble metals. They are often difficult to dissolve but deposit easily. The metals with low equilibrium potentials are active metals that dissolve easily but are more difficult to reduce.

Tab. 2 Standard electrode potentials of reactions of interest in metal electrolysis [1]

Reaction	*Standard electrode potential, E^o[V]*
$Au^+ + e^- \leftrightarrow Au$	1.692
$Au^+ + 3e^- \leftrightarrow Au$	1.498
$Cl_2 + 2e^- \leftrightarrow 2Cl^-$	1.395
$O_2 + 4H^+ + 4e^- \leftrightarrow 2H_2O$	1.228
$Pt^{2+} + 2e^- \leftrightarrow Pt$	1.188
$Pd^{2+} + 2e^- \leftrightarrow Pd$	0.987
$Hg_2^{2+} + 2e^- \leftrightarrow 2Hg$	0.788
$Ag^+ + e^- \leftrightarrow Ag$	0.799
$Cu^+ + e^- \leftrightarrow Cu$	0.520
$SbO_2^- + 4H^+ + 3e^- \leftrightarrow Sb + 2H_2O$	0.446
$AsO_2^- + 4H^+ + 3e^- \leftrightarrow As + 2H_2O$	0.429
$Cu^{2+} + 2e^- \leftrightarrow Cu$	0.337
$Bi^{3+} + 3e^- \leftrightarrow Bi$	0.215
$2H^+ + 2e^- \leftrightarrow H_2$	0.000
$Fe^{3+} + 3e^- \leftrightarrow Fe$	−0.037
$Pb^{2+} + 2e^- \leftrightarrow Pb$	−0.126
$Sn^{2+} + 2e^- \leftrightarrow Sn$	−0.136
$Ni^{2+} + 2e^- \leftrightarrow Ni$	−0.250
$Co^{2+} + 2e^- \leftrightarrow Co$	−0.277
$Cd^{2+} + 2e^- \leftrightarrow Cd$	−0.403
$Fe^{2+} + 2e^- \leftrightarrow Fe$	−0.440
$Zn^{2+} + 2e^- \leftrightarrow Zn$	−0.763
$Cr^{2+} + 2e^- \leftrightarrow Cr$	−0.913
$Mn^{2+} + 2e^- \leftrightarrow Mn$	−1.179
$Ti^{2+} + 2e^- \leftrightarrow Ti$	−1.630
$Al^{3+} + 3e^- \leftrightarrow Al$	−1.663
$Mg^{2+} + 2e^- \leftrightarrow Mg$	−2.363

Tab. 3 Electrolyte components of interest in electrowinning and electrorefining [2]

Component	*Role*
Metal ion	Source of metal
Anion	Maximize solubility
	Provide anode reaction
	Conductivity
	pH control
Auxiliary salt	Ionic strength and conductivity
	Buffering and pH control
Impurities	By-product for recovery
	Detrimental to efficiency or purity
Inorganic additives	Assist anode life
	Impurity removal
Organic additives	Cathode leveling
	Impurity control and improved efficiency
	Grain refiner
	Mist suppressor
Water	Solvent
Suspended solids	Incorporation into cathode or remove as cell sludge

An electrowinning or electrorefining electrolyte is not a simple aqueous solution of a metal salt. The electrolyte contains many compounds that all have specific functions. Some compounds have a positive effect, whereas others are detrimental. The electrolyte component types and their roles are listed in Table 3.

4.1.1.1 Basic Electrochemistry

The probability of an electrochemical reaction happening is evaluated as electrode potential. The reaction rate is measured as current density. Electrochemical reactions usually occur at the electrode surface at which conduction changes from ionic to electronic and charge transfers over the interface by an electrochemical reaction. The reactions that release electrons are anodic reactions. Reactions that consume electrons are cathodic reactions. The surface on which anodic reactions occur is an anode. Cathodic reactions occur on a cathode. Anodic reactions are oxidation or dissolution reactions, and cathodic reactions are reduction or deposition reactions. A reduction reaction consumes the electrons produced in an oxidation reaction. These are sometimes considered half-reactions of a redox reaction. The combination of oxidation and reduction reactions forms an electrochemical cell and the corresponding redox reaction is a cell reaction.

In the equilibrium state, the anodic and cathodic partial reactions of an electrochemical reaction have equal rates. The system is in a dynamic equilibrium state, and no net reaction occurs. For example, when a copper sheet is immersed in copper sulfate solution, in the equilibrium state the anodic dissolution rate of copper from sheet to solution equals the cathodic deposition rate from the solution to the surface of the sheet. Theoretically, one can calculate the equilibrium state of an electrochemical reaction from thermodynamic values. This is the standard electrode potential, E°, or equilibrium potential of the electrochemical reaction. The standard electrode potential corresponds to a determined standard state of 0.1 MPa, 25 °C, activity of reactive species of 1 or ideal solution of 1.0 mol L^{-1}, and equilibrium potential of any other state.

Thermodynamic calculations allow estimations on the tendency of a reaction to occur. If the change of the so-called Gibbs free energy, ΔG (kJ mol^{-1}), of the reaction has a negative value, the reaction can proceed in that direction. If the value of ΔG is positive, then the reaction cannot occur spontaneously. The more negative the ΔG value, the more likely is the reaction to occur spontaneously. A chemical reaction can proceed in either direction. The values for ΔG for the forward and reverse reactions are numerically the same but opposite in sign, since $\Delta G = \Delta G$ (products) – ΔG (reactants). The value of the free energy change of a chemical reaction changes only in sign when the direction of the reaction reverses. If the anodic reaction has a negative ΔG value, the corresponding cathodic reaction has a positive value. If the equilibrium potential of an electrochemical reaction is low, it will more likely proceed in the anodic direction (negative ΔG) than in the cathodic direction (positive ΔG). If the equilibrium potential is high, the reaction will more likely proceed in the cathodic direction.

The electrochemical series tabulates standard electrode potentials. Some sources call the electrochemical series oxidation/reduction potentials, electromotive series, and so on. The reference state of electrochemical series is the hydrogen evolution reaction, or H^+/H_2 reaction. Its standard electrode potential has been universally assigned as 0 V. This electrode is the standard hydrogen electrode (SHE) against which all others are compared. For example, the standard electrode potential of the Fe/Fe^{2+} reaction is −0.440 V and that of Cu/Cu^{2+} reaction is +0.337 V. The standard electrode potentials are calculated from Gibbs free energy values by Eq. (8) that is applicable only in the abovementioned standard state.

$$E^{\circ} = -\frac{\Delta G}{z \cdot F} \tag{8}$$

where ΔG is the Gibbs free energy change for the overall cell reaction and z is the number of equivalents involved in the reaction.

Under other conditions, the equilibrium potentials are calculated by the Nernst equation (9):

$$E = E^{\circ} - \frac{R \cdot T}{z \cdot F} \cdot \ln\left[\frac{c_R}{c_O}\right] \tag{9}$$

where

- R is the molar gas constant = 8.3143 J mol^{-1} K^{-1}
- T is absolute temperature (K)
- c_R and c_O refer to the concentrations of reduced and oxidized species, for example, solid metal and metal ion.

Instead of concentrations, one should strictly use activities in thermodynamical calculations. Most basic calculations assume that the system is ideal and activity of a reacting species equals its concentration. This is usually not valid. The difference is the activity coefficient, γ. The activity of a species is therefore its concentration multiplied by the activity coefficient.

Calculation of the equilibrium potentials of the electrochemical reactions allows estimation of the thermodynamical cell voltage of the electrolysis process. This is the theoretical minimum voltage that must be applied to the system to start electrochemical reactions. For example, in the electrorefining of copper the equilibrium potential of the anode reaction is the same as that of the cathode reaction. The thermodynamical cell voltage is zero. On the other hand, in the electrowinning of copper from sulfate electrolyte, the anodic reaction is oxygen evolution and cathodic reaction is copper deposition. The thermodynamical cell voltage is 0.89 V, which is the difference between oxygen evolution ($E^{o} = 1.228$ V) and copper deposition ($E^{o} = 0.337$ V). In addition to the thermodynamical cell voltage, polarization effects related to the kinetics of electrochemical reactions and ohmic effects related to the hardware increase the actual cell voltage needed to drive the cell.

4.1.1.1.1 **Polarization Effects** When the potential of an electrode changes, it is no longer in the equilibrium state. Increasing the potential starts an anodic net reaction, and decreasing the potential starts a cathodic net reaction. These generate a current flowing through the electrode. Consequently, forcing current through the electrode will change its potential. The deviation from equilibrium potential is polarization, and the difference between actual electrode potential and its equilibrium potential is overpotential. The signs of the overpotential and current are positive for anodic reactions and negative for cathodic reactions.

Increasing polarization increases the current and vice versa. If the polarization is low, that is, the potential shift per unit increase in current density is small, then no factors will retard the electrochemical reaction rate. If the polarization is high, the electrode potential must change considerably to obtain higher currents.

Polarization has various meanings and interpretations depending on the system under study. For an electrochemical reaction, this is the difference between actual electrode potential and reaction equilibrium potential. Anodic polarization is the shift of anode potential to the positive direction, and cathodic polarization is the shift of cathode potential to the negative direction. In an electrochemical production system driven with an external current source, polarization is a harmful phenomenon. It will increase the cell voltage and therefore production costs. A system that polarizes easily will not pass high currents even at high overpotentials. The reaction rates are therefore small.

Electrode polarization is not a simple phenomenon. Electrochemical reactions are heterogeneous reactions. They occur in several phases in subsequent steps. The reaction path of metal deposition usually includes the following steps:

- Transfer of reactive species from electrolyte to electrode surface
- Adsorption
- Charge-transfer step at the surface
- Surface diffusion
- Growth of the metal atom into the metal crystal structure.

Chemical reactions can happen before or after the charge-transfer step. Any step can be rate determining, that is, the slowest one determines the total reaction rate. As the electrode polarizes, the resulting overpotential consists of several factors. The most important ones are activation, concentration, and resistance overpotentials. The activation overpotential results from the limited rate of a charge-transfer step, concentration overpotential from the mass-transfer step, and resistance overpotential is the result of ohmic resistances such as solution resistance. Depending on the nature of the slowest step, the reaction is activation, mass transfer, or resistance controlled.

The relationship between overpotential and current density of a single, activation-controlled electrochemical reaction is the Butler–Volmer equation, Eq. (10). As the equation shows, the rates of anodic and cathodic partial reactions are exponentially dependent on the overpotential. The net current is the sum of the anodic and cathodic partial currents.

$$i = i_0 \cdot \left[e^{\frac{\alpha z \cdot F}{RT} \cdot \eta} - e^{\frac{-(1-\alpha) z \cdot F}{RT} \cdot \eta} \right] \tag{10}$$

where i is the current density and i_0 is the exchange current density corresponding to the reaction rate in the dynamic equilibrium. In Eq. (10), the factor α is the symmetry factor with values between 0 and 1. If α is higher than 0.5, the rate of the anodic partial reaction increases faster with increasing overpotential than the rate of the cathodic reaction. At values $0 < \alpha < 0.5$, the rate of the cathodic partial reaction increases faster. The high overpotential (and high current density) approximation of the Butler–Volmer equation is known as Tafel equation (11)

$$\eta = a + b \cdot \log(i) \tag{11}$$

In Eq. (11) η is the overpotential, the factor a describes the exchange current density and b is the Tafel slope (millivolts/decade).

Activation overpotential is related to the charge-transfer step. A reaction that has high exchange current density and low Tafel slope has low activation overpotential. One can visualize the activation overpotential as a threshold to be overcome before the reaction starts. An electrode reaction that has a low exchange current density also has high activation overpotential. It will not start easily. For example, the term ''hydrogen overvoltage'' describes the polarization required for hydrogen evolution to start. Some metals such as zinc have a high hydrogen overvoltage. The high hydrogen overvoltage of zinc means that hydrogen evolution on zinc does not begin easily and proceeds sluggishly.

Mass-transfer overpotential results from a finite mass-transfer rate from bulk electrolyte to electrode or vice versa. If the system is mass-transfer controlled, a limiting current density exists. The limiting current density is the maximum reaction rate under mass-transfer control. It increases as the concentration of the reacting species, their diffusion rate, temperature, or flow rate increase. In a system with limiting current density, the overpotential follows Eq. (12). The overpotential increases very rapidly when approaching the limiting current density.

$$\eta_d = \frac{R \cdot T}{z \cdot F} \cdot \ln\left(1 - \frac{i}{i_{\lim}}\right) \tag{12}$$

If the cathode potential is sufficiently negative, all of the electroactive species undergo reduction to metal as soon as they reach its surface. The current at which this point is reached is known as the *limiting*

current. A further decrease of the cathode potential results in depletion of metal ions at the surface. At this point, the transport of the ionic species through the Nernst boundary layer to the cathode surface becomes the rate-determining step; that is, the rate of metal reduction at the cathode has moved from kinetic control to mass transport control. To increase the rate of the electrowinning process under these conditions, it is necessary to increase either the ionic strength of the electrolyte or the mass-transfer rate to the surface.

The limiting current density is a function of the concentration of the depositing metal, its diffusion coefficient, and the thickness of the diffusion layer. A high-performance cell needs to have a high mass transport coefficient and enhanced electrode area in order to achieve a high limiting current I_{lim} (13):

$$I_{\mathrm{lim}} = k_{\mathrm{m}} \cdot z \cdot F \cdot A \cdot c \qquad (13)$$

In Eq. (13) k_{m} is the local mass transport coefficient D/δ, that is, diffusion coefficient divided by the thickness of diffusion layer. High mass transport rates can be achieved by electrode movement, including reciprocating movement of planar electrodes, moving bed electrodes, or rotating cylinder electrodes. The electrode movement will decrease the diffusion layer thickness, enabling higher limiting current density. A high surface area per unit volume can be achieved by the use of porous, three-dimensional electrodes.

For electrowinning and electrorefining of metals, mass transfer toward the cathode is of supreme importance. For example, in an electrorefining cell with vertical plate electrodes, there is a steady flow of electrolyte from one end of the cell to the other, but the flow velocity is small and the flow is not directed parallel to the electrode surface. It does not give significant mass transfer from anodes to cathodes. Natural convective flow, driven by the density difference caused by depletion of the solution at the cathode, the lifting effect of gas bubbles, and so on, is the most important factor. For example in copper electrorefining, depletion of copper from the cathode surface decreases the electrolyte density by several percent. Figure 2 shows an example of natural convection, the concentration gradients in electrorefining caused

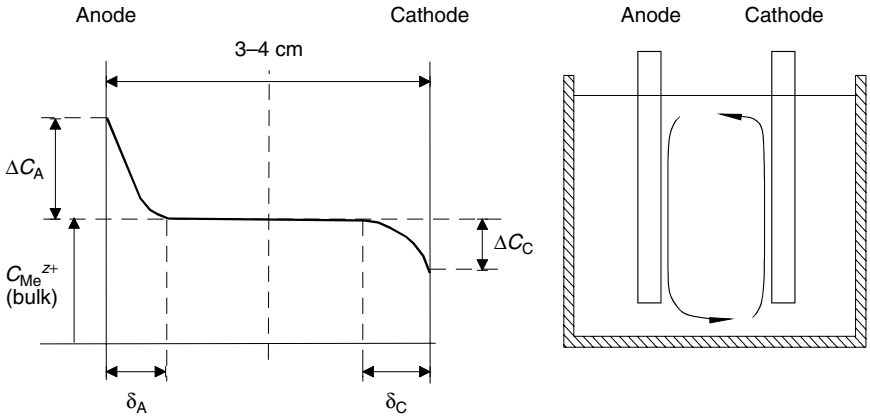


Fig. 2 The concentration gradients in electrorefining cell with vertical plate electrodes and the resulting natural convective flow pattern.

by anolyte of higher density and catholyte of lower density. In an electrowinning cell, the natural convection is in the opposite direction due to the lifting effect of gas bubbles at the anode.

The resistance overpotential does not directly influence the electrochemical reaction, but it decreases the charge transfer between anode and cathode. The resistance overpotential obeys Ohm's law, Eq. (14).

$$\eta_\Omega = I \cdot R_\Omega \tag{14}$$

As the solution resistance, current, or both increase, the effect of resistance overpotential becomes stronger.

The rate of an electrochemical reaction is measured as electric current. To allow comparisons between different systems, the rate is usually measured as current density. This is the current flowing through the electrode divided by the electrode area. A plot relating current density to electrode potential is a polarization curve. Current density usually follows a logarithmic scale. Figure 3 shows polarization curves for different overpotentials. The current densities are in linear and logarithmic scale. Different overpotentials have typical features. The activation overpotential of the Butler–Volmer equation, Eq. (10), increases sharply at low current densities. It has less effect at high current densities. The concentration overpotential, Eq. (12), has no effect at low current densities. As current density approaches the limiting current density, it approaches infinity. The resistance overpotential, Eq. (14), increases linearly with increasing current density to obey Ohm's law. The solution resistance causes an error called *IR drop* or *ohmic*

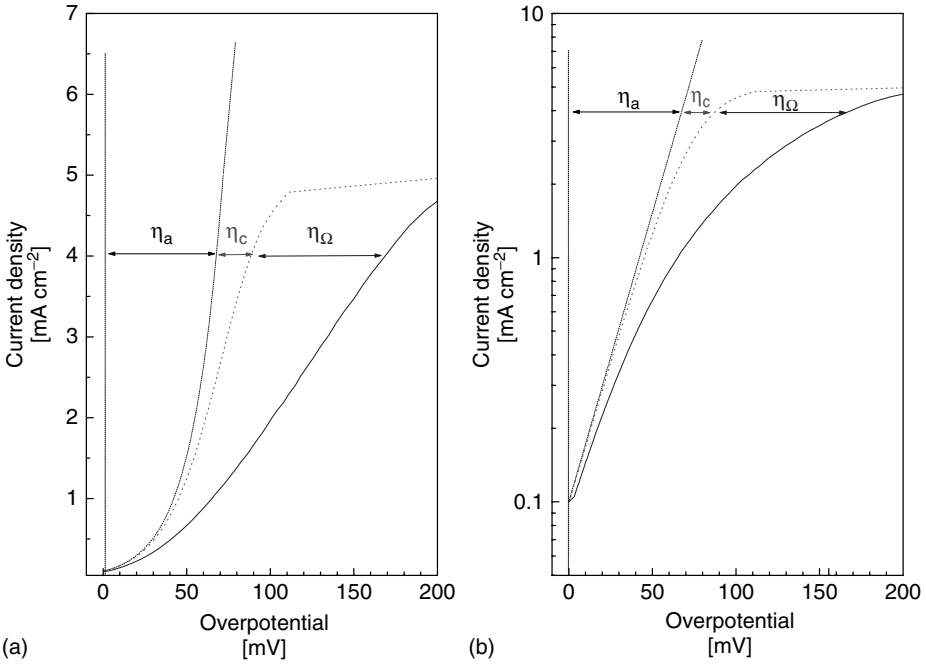


Fig. 3 Polarization curves for different overpotentials. In (a) the current density is on linear scale and in (b) on logarithmic scale.

drop. The *IR* drop distorts polarization curves such that for anodic curves the real electrode surface potential is always lower than the measured potential. For cathodic curves, the real surface potential is higher than the measured potential.

4.1.1.1.2 **Electrode Surface** The electrode surface is a complicated heterogeneous system as Fig. 4 shows. There is a net negative charge nearest the surface within a metal acting as a cathode. To maintain electroneutrality, there are adsorbed ions and water molecules next to the surface in the electrolyte side. Toward the bulk electrolyte there are cations surrounded by water molecules. This is the Helmholtz double layer. The inner Helmholtz layer contains water molecules and specifically adsorbed anions. The outer Helmholtz layer contains the second water molecule layer. From the Helmholtz double layer toward the bulk electrolyte are the diffusion layer and the hydrodynamic layer. In the diffusion layer, the concentration of species changes from that of the bulk electrolyte to that of the electrode surface. The diffusion layer does not move, but its thickness will decrease with increasing bulk electrolyte flow rate to allow higher reaction rates. The diffusion layer thickness is inversely proportional to the square root of the flow rate. The hydrodynamic layer or Prandtl layer has the same composition as the bulk electrolyte, but the flow of the electrolyte decreases from that of the bulk electrolyte to the stationary diffusion layer.

All electrochemical phenomena occur in these thin layers whose composition and properties differ from those of the bulk electrolyte. The inner Helmholtz layer is

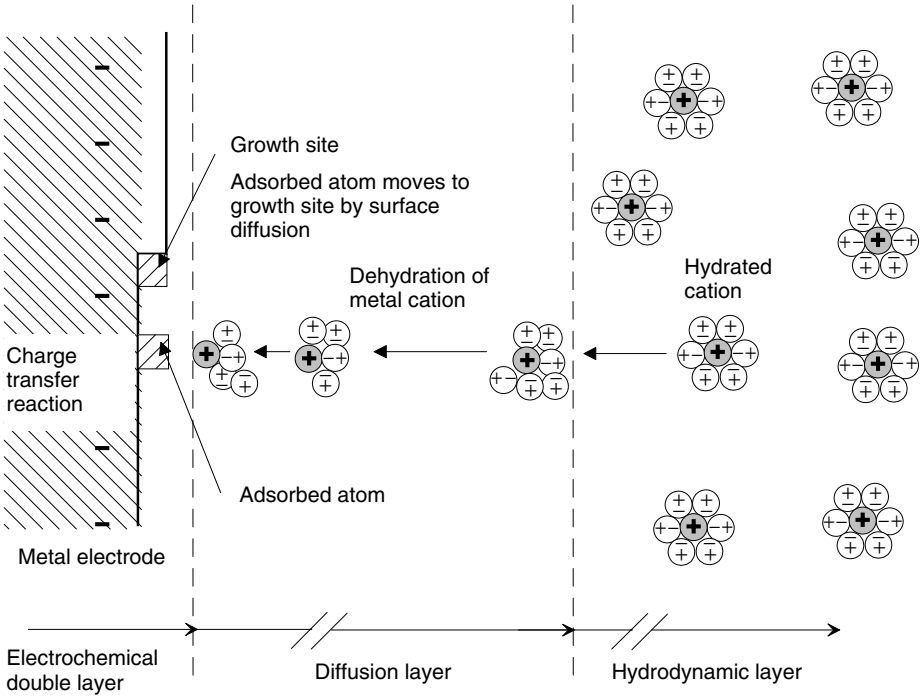


Fig. 4 Surface layers of an electrode during the cathodic deposition of metal.

about 1 nm and the outer one is about 4 nm thick. The diffusion layer is about 10–100 μm and the hydrodynamic layer is 0.1–1 mm thick. The double layer covers the entire metal surface and contains reactants and reaction products of all anodic and cathodic reactions.

The properties of the surface layers have a strong effect on the deposition process. The driving force of the electrochemical reaction is the potential difference over the electrochemical double layer. Adsorption of species can change this potential. For example, the additives used in electrodeposition adsorb in the Helmholtz layer. They can change the local potential difference, block active deposition sites, and so on. The thickness of the diffusion layer affects the mass-transfer rate to the electrode. The diffusion layer becomes thinner with increasing flow rate. When the diffusion layer is thicker than the electrode surface profile, local mass-transfer rates are not equal along the electrode surface. This means that under mass-transfer control, metal deposition on electrode surface peaks is faster than in the valleys and a rough deposit will result.

4.1.1.2 Current Distribution

The thickness distribution of electrodeposits depends on the current distribution over the cathode, which determines the local current density on the surface. The current distribution is determined by the geometrical characteristics of the electrodes and the cell, the polarization at the electrode surface, and the mass transfer in the electrolyte. The primary current distribution depends only on the current and resistance of the electrolyte on the path from anode to cathode. The reaction overpotential (activation overpotential) and the concentration overpotential (diffusion overpotential) are neglected. The secondary current distribution depends on both the resistance of the electrolyte path and the reaction overpotential, but the concentration overpotential is still neglected. Transportation of ions between the electrode and the solution through the diffusion layer determines the tertiary current distribution. All the three are involved in different degrees in the overall cathode deposition process.

Optimization of the electrochemical cell's geometry is the primary factor that determines the uniformity of current distribution. The two major geometrical arrangements are parallel-plate electrodes and concentric cylinders. Parallel-plate geometry is common in large-scale production of base metals when the metal concentration in the electrolyte is high. Cylindrical cells are used in the treatment of less concentrated solutions, in the recovery of noble metals, and also in the production of base metals. The current distribution between two parallel electrodes is only uniform when a nonconducting containment of the same cross section surrounds the interelectrode space. In all other cases, the current density is unevenly distributed, leading to higher current densities at the edge of the electrodes. Similarly, any protrusion on the surface has better access to current than the recesses, leading to faster deposition on the peaks of the microprofile (Fig. 5).

A secondary current distribution in the case of the parallel-plate electrodes brings the edge current density to a finite value. The edge current is still greater than that in the center of the electrode, unless the reactions are slow. The secondary current distribution can be estimated by using the dimensionless Wagner number (15)

$$Wa = \frac{\partial\eta/\partial I}{\rho \cdot L} \qquad (15)$$

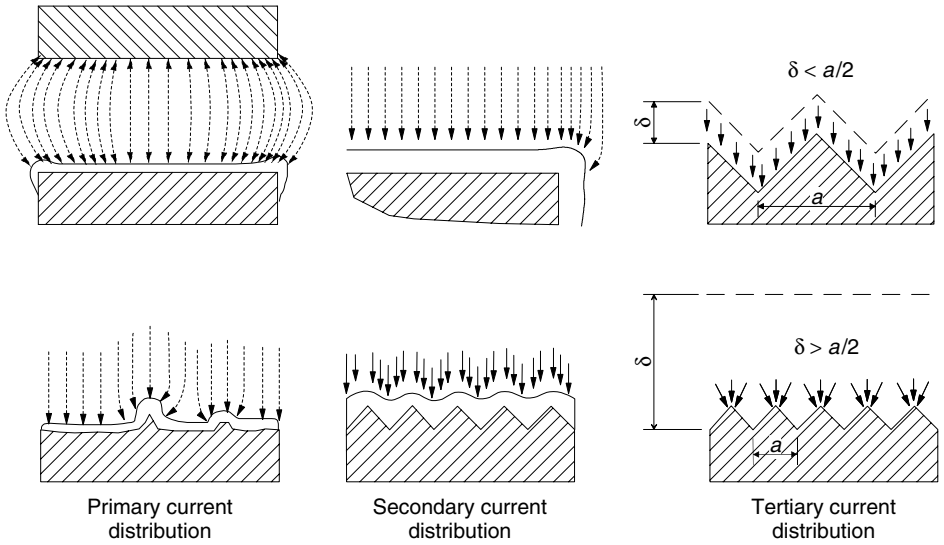


Fig. 5 The effect of primary, secondary, and tertiary current distribution on deposit thickness.

where $\partial\eta/\partial I$ is the reciprocal slope of a polarization curve (apparent charge-transfer resistance), ρ is the electrolyte resistivity, and L is a characteristic length. On the microprofile, the characteristic length can be, for example, the distance between two peaks. The Wagner number can be regarded as the ratio of polarization resistance per unit area to electrolyte resistance per unit area. When the Wagner number approaches zero, the primary current distribution prevails. The secondary current distribution is always more uniform than the primary current distribution and for $Wa \gg 1$, it becomes largely independent of cell geometry. Secondary current distribution predominates if the charge-transfer resistance is higher than the ohmic resistance. On the basis of the Wagner number, some general tendencies of secondary current distribution can be determined. The use of an inert electrolyte, such as sulfuric acid in electrowinning, will lower the resistivity and increase the effect of secondary current distribution. The use of organic additives to modify deposit growth can increase the polarization resistance and increase the effect of secondary current distribution. The lower the Tafel slope for the deposition reaction and the smaller the characteristic length, the more even will be the resulting deposit. The Tafel equation for charge-transfer overpotential (11) and the equation of ohmic overpotential (14) show that the charge-transfer resistance decreases logarithmically with current density and the electrolyte resistance linearly. For any process controlled by activation polarization (the charge-transfer step), increasing current density will lead to less uniform current distribution.

When the current density is so high that local concentration of the depositing metal changes, tertiary current distribution begins to affect the deposit thickness. The effect of concentration polarization depends on the ratio of surface concentration to bulk electrolyte concentration of the depositing metal. The surface concentration depends on the local mass-transfer

coefficient, that is, on the diffusion coefficient and diffusion layer thickness. The effect of tertiary current distribution depends on the characteristic length with respect to diffusion layer thickness. If the characteristic length (distance between two peaks) is larger than the diffusion layer thickness, the tertiary current distribution follows the surface profile. If the characteristic length is smaller than the diffusion layer thickness, the tertiary current distribution will decrease the smoothing effect of secondary current distribution. The diffusion length from the bulk electrolyte to the peaks is smaller, and, thus, more metal can deposit on the peaks, giving a more rough deposit [3, p. 126].

Figure 5 shows the effect of different current distributions on the potential distribution on the electrode surface. In primary current distribution, the current concentrates on protruding areas, because they are closer to the opposite electrode. Protruding areas also have lateral accessibility to current, and primary current distribution would give infinite current density on ideally sharp peaks. In secondary current distribution, the protruding areas polarize more owing to the higher local current density. This greater overpotential tends to oppose current flow and divert current to the recessed areas. The ohmic resistance to be overcome decreases with decreasing profile depth, and the equalizing effect is stronger on smaller profile depths. The current efficiency is often dependent on current density, and if the current efficiency decreases with increasing current density, it has a smoothing effect. If the current density is close to the limiting current density, the tertiary current distribution prevails. The higher the ratio of current to limiting current, the stronger will be the preferential deposition on microprofile peaks [4].

The electrodeposition process is very sensitive to current distribution because it is intrinsically unstable. The current concentration on sharp surface profile discontinuity leads to stronger growth on these areas of the cathode, including the edges. This latter can be eliminated by using a nonconductive mask on cathode edges or increasing the distance from anode to cathode edges. In parallel-plate cells, the latter can be done simply by using cathodes that are larger than anodes. If the cathode surface is not smooth, local current densities increase on whiskers and nodules, causing them to grow toward the anode.

4.1.1.3 Cell Voltage and Energy Use

The cell voltage depends on the thermodynamic equilibrium potentials of the anode and cathode reactions, kinetic overvoltages, and ohmic resistances caused by the cell hardware. The cell voltage can be calculated using Eq. (16)

$$E_{\text{cell}} = E_{\text{cathode}} - E_{\text{anode}} - \Sigma(\eta) - i \cdot R_{\text{cell}} \tag{16}$$

In Eq. (16), E_{cathode} and E_{anode} are the equilibrium potentials of the electrode reactions, $\Sigma(\eta)$ is the sum of all polarization effects on the electrodes and $i \cdot R_{\text{cell}}$ is the ohmic loss due to cell construction. As the electrolysis systems consume energy, the cell voltage sign is negative.

It is important to minimize the overpotential and ohmic components of cell voltage in order to maximize energy efficiency. This can be done by using conductive, catalytic electrodes and membranes, small interelectrode (or membrane-electrode) gaps, and by careful choice of the counterelectrode chemistry to minimize the equilibrium cell voltage. Costs at the optimum usually become higher in the case of more dilute solutions as the mass-transfer-imposed limiting current is lower.

The electrolytic separation of metals from waste solutions is therefore problematic as these solutions are usually dilute in metal ion.

The metal deposition rate from aqueous solutions can be expressed as a current density (A m^{-2} of electrode surface) as the rate is linked to the current density by Faraday's law (1). In modern copper electrorefining and electrowinning, the current density is 300–350 A m^{-2}, while in zinc operations it is greater, 450–500 A m^{-2} or more. According to Faraday's law, enormous amounts of electricity are needed for the tens or hundreds of kilotons of metal produced in a typical electrolysis plant per year.

The energy costs are related closely to the cell voltage. The energy consumption can be estimated by specific energy consumption w, Eq. (17)

$$w(J\ g^{-1}) = \frac{z \cdot F \cdot E_{\text{cell}}}{\eta \cdot M} \qquad (17)$$

where E_{cell} is the cell voltage (V). The specific energy consumption for a particular metal and process is calculated using the electrochemical equivalent of the metal, the voltage needed to drive the process, and the current efficiency. The common way to express energy consumption is to use kWh kg^{-1} by dividing the right-hand side of Eq. (16) by 3600. To lower the specific energy consumption, it is necessary to lower the cell voltage and/or increase the current efficiency.

Energy costs are an important component of any electrolysis process. The energy consumption depends on the operating current density and cell voltage. The operating current densities are well below the limiting current density set by the mass-transfer rate. It has been found that operation at some fraction of the limiting current density is necessary to avoid excessive cell voltage and operating problems such as uneven deposit growth or low current efficiency. The optimum current density is also a compromise between reduced capital costs and higher operating costs as current density is increased.

Examples of the current density–cell voltage dependence are shown in Figs 6 and 7. In Fig. 6, the system is the electrorefining of copper. The thermodynamical cell voltage is zero. When the cell is operated, polarization effects determine the cell voltage. Increasing the current density increases polarization. Upon increasing cathode polarization, the deposit morphology changes to more rough and finally to powder when limiting current density is reached. The anode polarizes continuously unless anode impurities cause anode passivation, after which the anode potential increases sharply. In Fig. 7, the system is the electrowinning of copper. The thermodynamical cell voltage is 0.89 V. The total cell voltage depends again on the electrode polarization, which depends on the operating current density. The effect of cathode polarization is the same as in electrorefining. The anode polarizes continuously as the anode reaction is the decomposition of water, Eq. (4).

Table 4 shows examples of the energy consumption of some electrorefining and electrowinning processes [5, p. 163].

4.1.1.4 Factors Affecting Deposit Quality

A metal deposit consists of crystals, which originate as nucleus. Formation of a new crystal begins with the adsorption of a single ion from the electrolyte reducing to an adatom. This adatom is loosely bound to the surface and at any given moment, a certain number of adatoms are present on the surface. Adatoms can dissolve again, form a new nucleus, or grow into an existing crystal. The adatoms move around

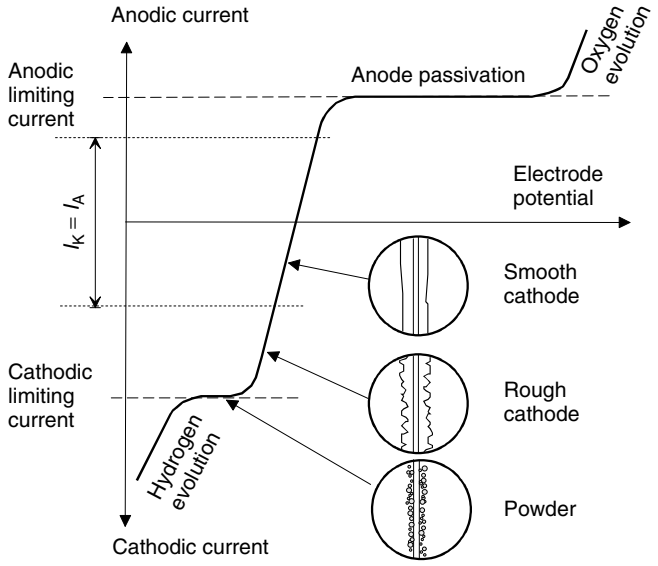


Fig. 6 Current density–cell voltage dependence for the electrorefining of copper.

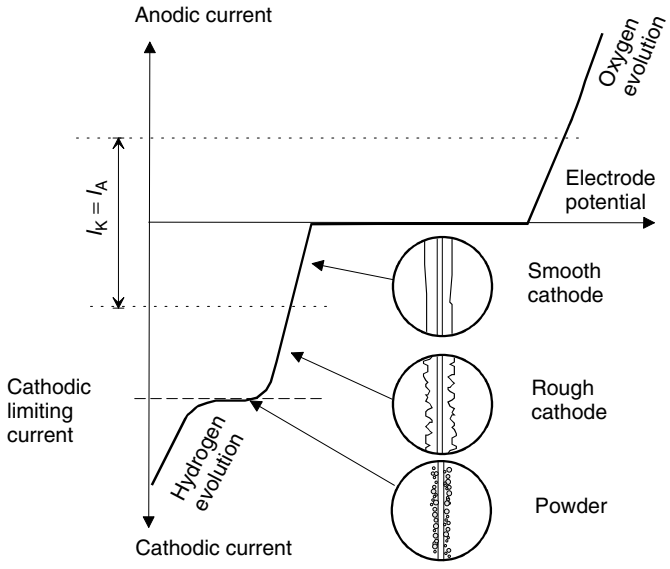


Fig. 7 Current density–cell voltage dependence for the electrowinning of copper.

by surface diffusion. If surface diffusion is hindered or adsorbed compounds block the growth sites, the adatom concentration builds up. This will lead to formation of new nuclei and, thereby, new growth sites. When adatoms meet, they become more

Tab. 4 Comparison of energy use in some electrorefining and electrowinning operations [5, p. 163]

Metal	*Electrolyte*	*Temperature [°C]*	*Thermodynamical voltage [V]*	*Cell voltage [V]*	*Current density [A m^{-2}]*	*Current efficiency [%]*	*Energy consumption [kWh kg^{-1}]*
Electrorefining							
Ag	Nitrate	25	0	2.0–2.8	300–400	93–98	0.7
Cu	Sulfate	60	0	0.2–0.3	150–350	92–98	0.2–0.3
Ni	Sulfate/chloride	60	0	1.9	200	98	2.0
Pb	Fluosilicic	40	0	0.45	200	90–92	0.2
Electrowinning							
Cu	Sulfate	40	0.90	1.9–2.0	170	85–90	1.8–2.0
Co	Sulfate	60	1.52	3.5–4.0	160	80–90	3.7
Zn	Sulfate	35	2.0	3.6	400–1000	90–94	3.2

stable and larger clusters of adatoms may form. The larger the cluster of adatoms, the more stable the cluster becomes. If the cluster reaches a certain critical size, it will not dissolve anymore, and will become a nucleus.

Nucleation has a strong influence on the structure of the deposit. The nucleation rate depends on the number of adatoms on the surface. This adatom concentration will be higher for higher overpotentials resulting in a higher nucleation rate. Nucleation can be instantaneous or progressive. When instantaneous nucleation takes place, the number of nuclei is independent of time. When progressive nucleation takes place, the number of nuclei increases with time. The crystal structure resulting from an electrodeposition process depends on the relative rates of nucleus formation and the growth of existing crystals. Fine-grained deposits result from conditions that favor nucleus formation, while larger crystals are obtained in conditions that favor growth of existing crystals. Generally, a smaller crystal size is obtained with increase of the cathode polarization. Fischer uses the term ''inhibition'' to describe retardation of the deposition process [6]. Inhibition is due to the presence of substances other than the depositing metal at the surface of the electrode, in the double layer, or in the diffusion layer. Inhibiting compounds change the polarization effects at the surface.

Fischer has developed a classification system for metal deposits, Fig. 8. The five basic types of morphologies are as follows:

- Field-orientated isolated (FI) crystals
- Base-oriented reproduction (BR)
- Twinning intermediate (Z)
- Field-oriented texture (FT)
- Unoriented dispersion (UD).

The grain refinement is achieved by an increase in the nucleation rate. The first four deposit types are seen with two-dimensional nucleation, but UD type requires three-dimensional or ''secondary'' nucleation. The nucleation intensity increases with increasing inhibition and deposition current density. The FI type of crystals is usually seen at low inhibition. With increasing current density, the deposit form changes from whiskers to prismatic crystals, dendrites, and finally powder deposits. The BR type is seen at moderate inhibition level and the deposit follows the structure of the base material.

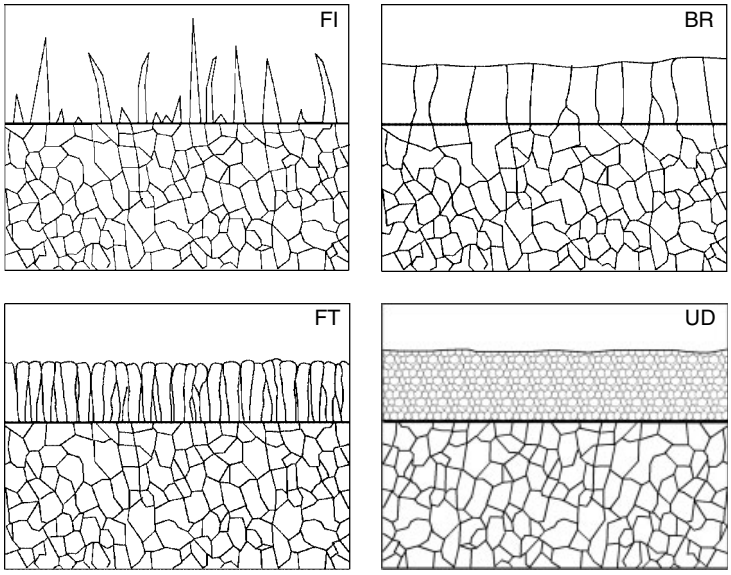


Fig. 8 The basic deposition types according to Fischer [7].

The deposit has sufficient time for lateral growth. With time, the surface roughness of BR deposit can increase, resulting in FI type deposit. The Z type is an intermediate between BR and FT types. The FT type consists of a large number of elongated crystals perpendicular to the substrate. The FT type is seen at relatively strong inhibition and it forms a coherent deposit. The UD type is a coherent deposit of a large number of small crystals. The UD type is seen at the strongest inhibition, but if the current density is too high, the deposit will be in powder form. The organic additives used in electrodeposition function as inhibitors, promote lateral growth, and lead to dense deposits.

The preferred types are BR type, twinning intermediate type (Z), FT type, and FT mixed with UD type. The best deposits have often been obtained at the border of the FT–UD transition. FI crystal type growth can lead to dendritic growth at high current densities and to prevent this, a higher inhibitor concentration is necessary. UD type can lead to deposition of powder, which in turn can start formation of nodular deposit. To prevent powder formation, the current density and inhibitor concentration should not be too high. The deposit is considered poor quality when it is rough and nodular. Nodules often cause short circuits in the electrolysis tank with a reduction in current efficiency and may cause entrapment of anode slimes, reducing cathode purity.

Winand has further developed a theory of electrocrystallization of polycrystalline metal deposits. This theory is summarized in the Winand diagram, Fig. 9, which uses the degree of inhibition and the ratio of operating current to limiting current to define areas for different deposit types. The degree of inhibition is related to the charge-transfer overpotential and the ratio of operating current to limiting current to the concentration (or diffusion) overpotential [8].

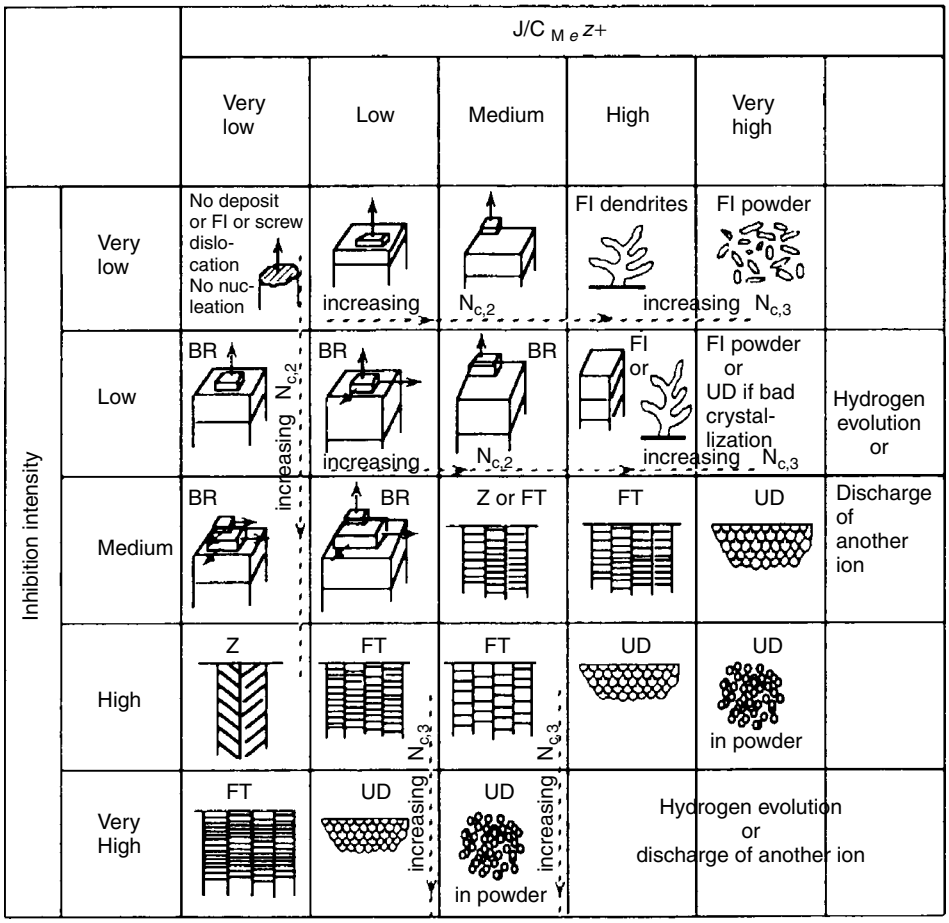


Fig. 9 Winand diagram showing different possible types of polycrystalline deposits as a function of i/i_{lim} and inhibition intensity [8].

Activation-limited growth tends to favor compact columnar or equiaxed deposits, while mass transport–limited growth favors formation of loose dendritic deposits. The deposit morphology is modified by using additives. Additives act as grain refiners and levelers because of their effects on electrode kinetics and the structure of the electrical double layer at the cathode surface. Additives that reduce primarily the nucleation overpotential can be considered to be grain-refining additives because of increased secondary nucleation events. Additives that increase the deposition overpotential at a given current density, for instance, by altering the Tafel constants, can be considered deposit-leveling additives. Since additives are typically present in very small concentrations, their transport to the electrode is nearly always under diffusion control and sensitive to flow variations.

There are several factors affecting cathode purity that are related to the additives. During the electrolysis process, incorrect additive levels can cause excessive edge growth, nodular growth, or dendrites.

These can cause short circuits or entrapment of the electrolyte. Entrapment of electrolyte in the deposit, with or without enrichment of additives, will lower the cathode quality. Codeposition of electrolyte impurities can lead to cathode contamination. A nodule forms when a conductive or semiconductive particle remains on the cathode surface. The deposit will then grow around the particle encasing it. The nodular growth can be initiated by several factors. The nodule becomes larger with the continuation of electrolysis and more contaminated with slime particles. Nodules can also lead to short circuits. The local current efficiency can vary depending on the inhibitors and local current density. This can also result in rough deposits, if the current efficiency increases with current density.

4.1.2 Electrolysis Technology

The design and construction of an electrolysis cell depend on the cell reactions, metal concentration, and production rate. Most of the cell designs operate continuously and at steady state. In electrorefining operations, the electrolyte composition does not change much. In electrowinning operations, the metal concentration decreases and acidity often increases. Most of the cell design parameters deal with mass transport of the reacting species and maintaining the purity and composition of the electrolyte.

4.1.2.1 Cell Construction

The traditional electrolytic cells are simple, rectangular, open tanks with parallel-plate electrodes. Cells are generally made from concrete and lined with lead or polyvinylchloride (PVC). The traditional configuration is lead-lined concrete. Concrete cells with plastic liners were popular in the 1970s and 1980s. The corrosion-resistant lining is necessary to protect the tank from the acid corrosive electrolyte. Polymer-modified concrete is nowadays used to avoid the lining of cells. The polymer concrete cells became popular in the late 1980s and 1990s in copper production with conversions to permanent cathode technology and in new tankhouses. Polymer concrete is a corrosion-resistant construction material that withstands the mechanical wear in the electrode handling.

In the traditional cells, the cathodes and anodes are flat plates that are placed parallel to each other. Anodes and cathodes are suspended alternately in the cells with precise spacing. In some cases, it is necessary to separate anode and cathode electrolytes. This is done by using separators, with diaphragms that allow solution flow or membranes that allow only anion or cation transfer. The deposition of metal onto the cathodes is a batch process. The cathodes are removed from the cell when the deposit has grown thick enough. The flow rate of the electrolyte is kept low so that possible solids fall to the bottom of the cell and do not contaminate the cathodes. Figure 10 shows an example of the traditional parallel-plate cell.

Diaphragm or membrane cells are used in special cases where impurities are easily codeposited on the cathode, anode slimes are a problem, or the anolyte and catholyte have a different composition. A diaphragm cell can be made by enclosing either the anodes or the cathodes in a rigid bag. Enclosing anodes in a bag is an efficient method to collect anode slimes, but the anode compartments must be cleaned quite often. Enclosing cathodes in a bag requires purification of the anolyte in a

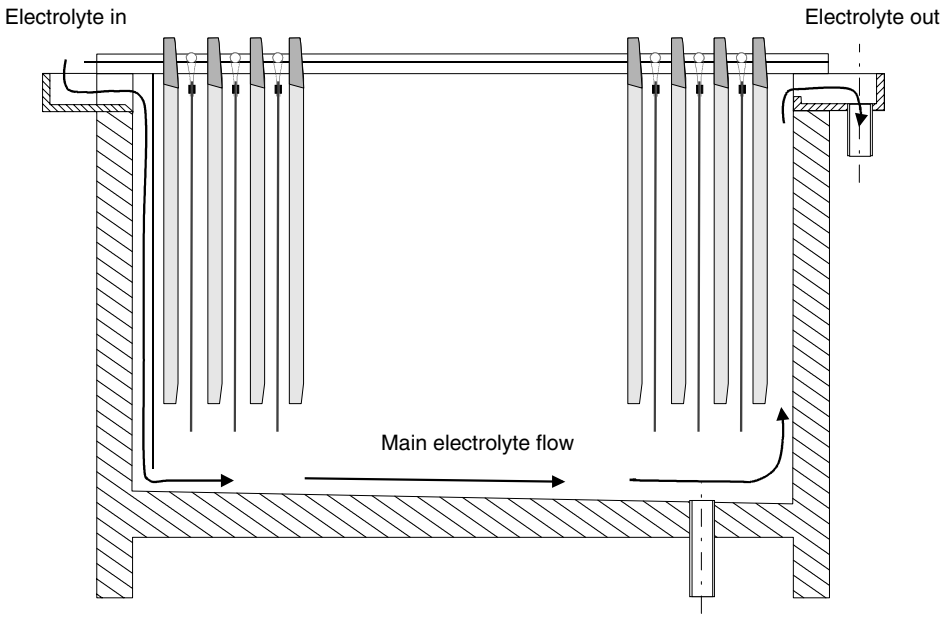


Fig. 10 Parallel-plate cell used in traditional electrowinning and electrorefining.

separate circuit before pumping it to the cathodes. Diaphragm cells are used, for example, in nickel production, noble metal electrorefining, and in cells with chlorine evolution as anode reaction.

With open top tanks, there are difficulties for plant operators in the form of acid mist. The acid mist is generated when gas bubbles formed at electrodes eject electrolyte droplets into the air above the cells. The gas evolved is usually oxygen at the anode or hydrogen at the cathode. The fine acid mist is an occupational risk and damages the building structure. Increased ventilation rates become costly, cause significant electrolyte temperature loss, and merely transfer the contamination outside the building. Prevention of acid mist can be done using physical means, such as floating balls or surface-active chemicals.

The space–time yield of conventional planar plate tanks is low, of the order of 70 kg of metal per cubic meter and day [9]. When the metal concentration of the electrolyte decreases, the parallel-plate geometry is no more efficient. To increase the deposition rate, there are two possibilities. Increasing the flow rate will increase the mass-transfer rate and the limiting current density as the diffusion layer becomes thinner, see Eq. (13). Increasing the effective cathode area will increase the space–time yield. It is common for process liquors and rinse waters to contain relatively low concentrations of dissolved metal (typically <1000 ppm). Such low levels require high mass transport to the cathode surface and/or high–surface area electrodes to provide a high production rate of metal with reasonable current efficiency [10].

Three types of mass transport occur in an electrochemical system. Diffusion is the result of a concentration gradient. Convection is the result of bulk motion of the

electrolyte, and it can be natural or forced. Migration is the movement of charged particles caused by an electric gradient. The usual methods to improve mass transfer are to use forced electrolyte flow or moving cathodes. The forced convection can be done, for example, by pumping electrolyte or using gas sparging under electrodes. Rotating cylinder electrodes (RCE) have been used for many years in specialized reactors for metal removal. The rotating cathode is usually surrounded by inert anodes, and it rotates at speeds of 50–300 rpm. Specialized, divided, and scraped RCE reactors can be used for the continuous removal of metal as powder [10], Fig. 11.

Fluidized-bed and packed-bed reactors have increased surface areas as compared to plate electrodes. A packed bed is made of stationary metal particles through which the electrolyte is pumped. The difficulty is that electrodeposited metal soon plugs the bed. A fluidized-bed reactor consists of a particulate electrode material in the full electrolyte volume of ascending liquid flow. An intermediate between the fluidized bed of particles and a packed bed is the spouted bed. In spouted bed cells, only the spout fraction of the bed is in vigorous motion and the rest of the particles descend slowly to the spout where they are again swept rapidly up to the top of the bed and fall down. The bed of particles that is agitated by the upward flowing electrolyte is held in shape by using a diaphragm. The current is fed to the bed by using a current feeder rod or mesh. In the fluidized bed, metal particles can be grown from a seed less than a millimeter to few millimeters. The calculated current density per the diaphragm area (''superficial current density'') can be 1–2 magnitudes higher than that of a plate electrode.

The fluidized-bed electrode was invented in the late 1960s. Three major problems prevented their use in industrial metal electrowinning: high cell voltages, adherence of bed particles to current feeder, and adherence and penetration of the diaphragm by bed particles [9]. The fluidized beds had also problems in maintaining all the particles in motion throughout the cell and the energy cost associated with pumping the electrolyte to achieve the upward solution flow.

Porous electrodes have the same advantages as particle bed reactors, but they have the difficulty of becoming plugged as the metal deposit builds up. Sometimes this is not important if the value of the deposited metal is high and the electrode may be sacrificed. An example of such a system is the Zadra cell used in gold electrowinning, where the cathode is a steel mat that can be melted away when gold is recovered.

Many commercial cell types can be used for metal recovery. The cells are based on different concepts, and the process objectives may be recovery of valuable metal, metal removal from the effluent to meet discharge requirements, and so on. Some of the developed and commercialized cell types include the following [3]:

- Chemelec cell with fluidized-bed glass beads within electrodes. The cell performance is achieved mainly by electrolyte movement. The cathode product is removed by scraping or used as anodes.
- Eco-Cell with cylindrical rotating cathode. The cell performance is achieved by electrode movement and high electroactive area. The cathode product is removed by scraping or fluidization as powder.
- ER cell with packed bed of carbon particles within parallel-plate electrodes. The cell performance is achieved by electrolyte movement and high electroactive

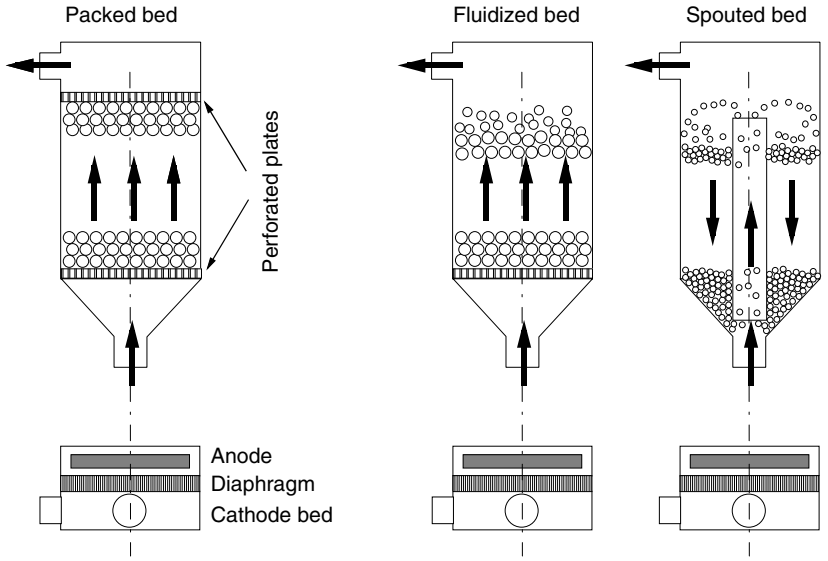


Fig. 11 Packed bed, fluidized bed, and spouted bed.

area. The cathode product is removed by discontinuous leaching.

- Goecomet reactor with rotating tubular bed or impact rod cathode. The cell performance is achieved by electrode movement and high electroactive area. The cathode product is removed continuously as sludge.
- Retec cell with vertical metal or carbon foam cathodes. The cell performance is achieved mainly by high electroactive area. The cathode product is removed by dissolving or is used as anodes.
- The Swiss-Roll cell with a sandwich construction made from a pair of thin, flexible electrodes, wound with the interposition of a pair of separators for electrical insulation. The cell performance is achieved mainly by high electroactive area. The cathode product is removed by replacing the electrode module.

4.1.2.2 **Current Feed**

The electrochemical cells use direct current in either monopolar or bipolar connection. In a monopolar cell, there is an external electrical connection to each electrode. The cell voltage is applied between each anode and cathode. In a monopolar cell, the anodes and cathodes alternate. Both faces of the electrodes are active. Monopolar cells require low voltage, but high current. Bipolar cells use only two electric connections to the end electrodes. In a bipolar cell excluding the end electrodes, the electrodes have one anodic and one cathodic side. The cell reaction proceeds when there is large enough potential difference between two electrodes, Fig. 12.

Most of the electrorefining and electrowinning processes use the monopolar connection. The overall reaction chemistry for electrowinning with bipolar electrode is similar to that of the monopolar process. The bipolar connection is difficult to construct because of problems in metal deposit harvesting. The bipolar cell design results in significantly increased cell voltages. This can result in power losses through bypass and stray currents in the

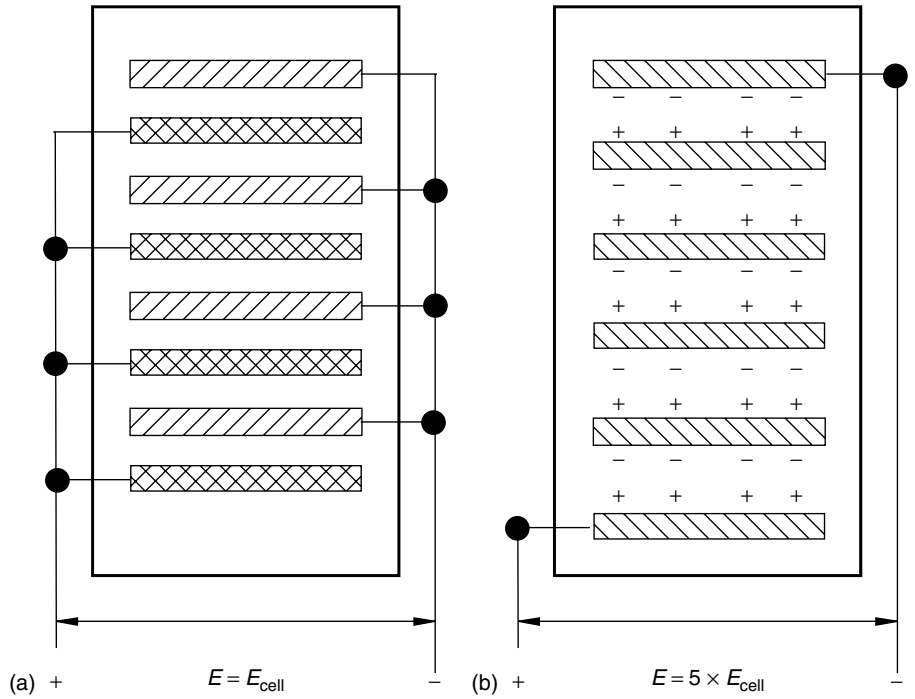


Fig. 12 Monopolar (a) and bipolar (b) cell connections.

cell and in the tankhouse and can affect occupational safety.

In a typical electrowinning or electrorefining operation, the electrolytic cells are rectangular tanks, and each contains 20–50 electrode pairs. An additional anode is required in each cell to ensure deposition on both sides of each cathode. The electrodes are designed to rest on busbars that supply electricity. The busbars are situated outside the top of each tank, one for the anodes and another for the cathodes. Thus, the electrode pairs in each tank operate in parallel. In the tankhouse, the cells are arranged in sections with the banks of cells connected in series and parallel to obtain optimum use of the electrical power, while keeping the voltage to earth at any point at a low level that is not a risk to the personnel.

In the traditional parallel-plate cells, the Walker system is the most commonly used electrical arrangement (Fig. 13). In this system, the current flows from a copper busbar on one side of the cell to the anodes, and the cathodes are connected to another busbar on the opposing side of the cell. The second busbar feeds current to the anodes of the second cell, and so on. In the Walker system, only one side of each electrode is connected to the electric circuit. The intercell busbars do not require so much thickness as the end busbars as the current flows through the path of least resistance.

The Walker configuration is prone to current imbalance caused by contact resistances. The current will flow through the least resistant path, so every anode–cathode pair will have a unique

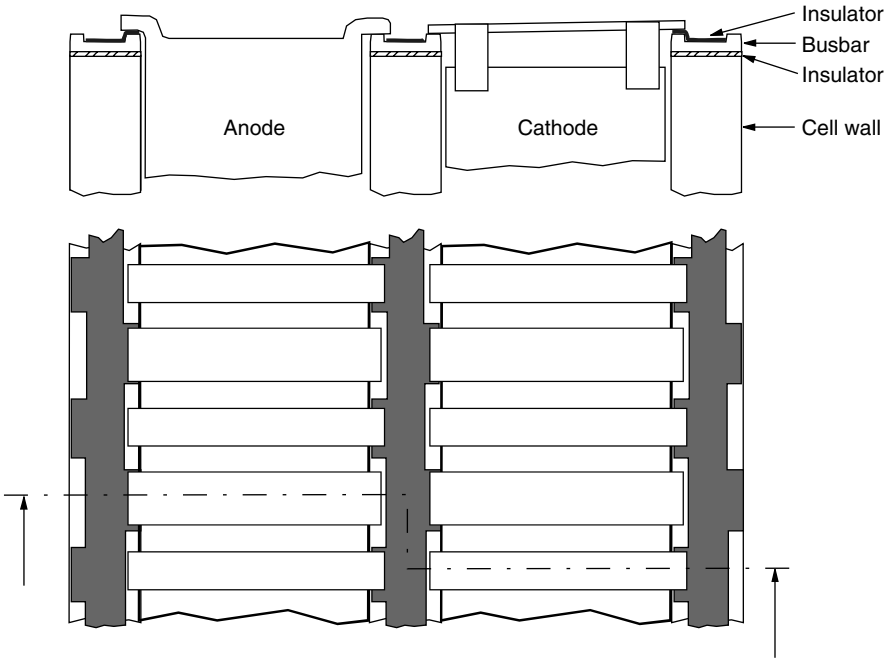


Fig. 13 The Walker system current feed to electrolysis cells.

voltage. Highly different voltages will lead to uneven current distribution in the cell. Methods to overcome the electrical problems of Walker configuration were developed a long time ago. The first attempt was the Whitehead system. In the Whitehead system, each cathode bar rests upon the supporting lug of the corresponding next cell anode. This system halves the number of connections in the tankhouse, but it was not taken into use because electrodes resting upon one another made the operation cumbersome. Later developments have focused on the intercell busbar arrangement by improving contacts and adding more current-distributing bars.

Usually, the current density applied in the electrolysis process is restricted by the limiting current density. As the current density increases too close to the limiting current density, serious deterioration occurs in the cathode quality. Pulsating current (PC) and periodic current reversal (PCR) are two techniques that permit an increase in current density without lowering the cathode quality. Early studies on copper, gold, and cadmium suggested that PC leads to refinement of the grain structure, which was attributed to the high instantaneous current densities favoring nucleation. Later studies indicated that nonsteady-state and steady-state mass transport conditions are the most crucial parameters. The deposit structure in PC depends on the nucleation mechanism and adsorption phenomena as well as on mass transport conditions, but a general theory in this field is still lacking [11].

PCR is a method by which a forward current is applied for a length of time followed by a quick current reversal. The interruption of the current flow in PC and

the change of polarity in PCR decrease the concentration gradients at the electrode-electrolyte interface. The effective current density with PC and PCR are equal to (18) and (19):

$$i_{\text{eff}} = \frac{i_+ \cdot t_+}{t_+ + t_0} \tag{18}$$

$$i_{\text{eff}} = \frac{i_+ \cdot t_+ - i_- \cdot t_-}{t_+ + t_-} \tag{19}$$

where i_{eff} is effective current density, i_+ and i_- are the current densities of the deposition and reverse phases, t_+ is the duration of the deposition phase, t_0 is the duration of the dead phase in PC, and t_- is the duration of the reverse phase in PCR.

In PCR, the forward to reverse period ratio is typically between 20/1 and 30/1. The PCR has two main effects. The current density may be increased without anode passivation. The reversal current depletes built up metal concentration within the anodic boundary layer. This helps in avoiding the precipitation of metal salts, which is one of the causes of anode passivation. On the cathode, thinning of the diffusion layer and selective removal of nodules during the reverse current phase result in smoother deposits. The major disadvantage of PCR is higher energy costs.

4.1.2.3 Permanent Anodes

Permanent anodes are used in the electrowinning of base metals such as copper, nickel, cobalt, and zinc. The only function of the anode is to transfer electrons from the electrolyte to the external circuit. The usual reactions on the anode surface are oxygen evolution in sulfate electrolytes and chlorine evolution in chloride electrolytes. The permanent anodes have the disadvantages of maintenance costs to clean the anode surfaces and restore the active area, higher power consumption as compared to electrorefining, and low cathode quality if it becomes contaminated owing to the deterioration of the anodes.

The traditional anodes for base metal electrowinning from sulfate solutions have been lead alloys containing calcium–tin, silver, or similar combinations. One potential concern with the lead anodes is the incorporation of lead corrosion products lowering the cathode purity. The other factor that is important to process economics is the high oxygen evolution overpotential that increases the cell voltage. More recently, anodes of titanium coated with precious metals, metallic oxides, and/or their alloys are manufactured. These types of electrodes are referred to as *dimensionally stable anodes* (*DSAs*) or ''inert'' anodes. The advantages of the DSA-type anodes include higher purity deposit, lower anode overpotential, and high corrosion resistance.

In sulfuric acid solutions, the main anodic reaction is oxygen evolution by the decomposition of water. In these solutions, the traditional anode materials have been lead alloys, mainly due to their low cost. Lead–silver alloy is commonly used in traditional zinc electrowinning operations. Antimonial lead is used in copper winning, but new high-purity processes have led to a change to lead–calcium–tin alloys. Metal silicide anodes (Chilex anodes) were formerly used for copper but are still used in some cobalt electrowinning operations. Nickel electrowinning anodes are often made from pure lead [12]. Proprietary anode compositions, anode treatments, and operational practices have been developed to overcome the problems of lead anodes. RSR Technologies has developed a rolled Pb–Ag–Ca–alloy with controlled surface grain structure for zinc electrowinning.

The controlled grain structure allows forming an adherent protective oxide coating within days when placed in an electrowinning cell without the necessity of being preconditioned [13]. In sulfuric acid electrolytes, the introduction of cobalt ions into the solution reduces the anode overvoltage and the corrosion rate of lead.

The electrocatalytically active layer of lead dioxide has a high overpotential toward oxygen evolution, and this causes a significant increase in energy use. The formation of lead dioxide layer happens at potentials of 1.6 V. This means an additional overvoltage of 0.4 V as compared to the equilibrium potential of oxygen evolution. The formation of lead dioxide requires corrosion of lead, which in sulfate solutions results first in the formation of lead sulfate. Crystallization of lead sulfate increases the anode potential to such high values that formation of lead dioxide becomes possible. During operation, it is important to keep a PbO_2 electrode well above the PbO_2-PbSO_4 potential; otherwise, the corrosion of the anode will occur. Despite the formation of a passive film of lead dioxide, there is generally a slow continued corrosion, which leads to the thickening of the PbO_2 film.

The DSA-type anodes are ''inert'', coated anodes made of a valve metal (titanium, niobium, or tantalum) base coated with an electrochemically active coating. The active coating is made either of noble metals or of mixed metal oxides. Noble metals in active coatings are usually platinum or platinum alloys. Mixed metal-oxide coatings contain active oxides and inert oxides; the active components are usually ruthenium dioxide (RuO_2) and iridium dioxide (IrO_2) and the inert components are mostly titanium dioxide (TiO_2) and other oxides such as tantalum oxide (Ta_2O_5). Oxide coatings are alloyed with inert oxides in order to improve the stability of the oxide layer and to decrease the cost. The first application of ruthenium oxide–coated anodes was the replacement of massive graphite anodes in chlor-alkali cells at the end of the 1960s *cf.* [14, 15]. Further developments have increased the use of oxide anodes to other electrochemical processes, environmental engineering, electroplating, and cathodic protection. $IrO_2 + Ta_2O_5$ is the preferred catalytic coating for O_2 evolving anodes in acidic sulfate electrolytes.

ELTECH System Corporation has developed and patented a mesh-on-lead (MOL) anode for primary copper electrowinning operations. The MOL anodes have performance advantages relative to standard Pb–Ca–Sn anodes, including lower oxygen evolution overpotential, improved cathode quality, and minimized lead sludge generation. The addition of cobalt is eliminated because of stabilized lead substrate, and current efficiency is improved because of reduced short circuits [16].

4.1.3 Copper Production

World copper refinery production in 2002 was 15 500 000 tons. Primary copper production was 13 780 000 tons, of which 11 100 000 tons (80%) was electrorefined and 2 680 000 tons (20%) was electrowon [17]. Pyrometallurgical processes employ high-temperature chemical reactions to extract copper from its compounds. These processes are generally used with copper sulfides and high-grade oxides. The use of high temperatures has several advantages: chemical reaction rates are rapid, some reactions that are impossible at low temperature become spontaneous at higher

temperature, and liquid phase facilitates separation of the metal from the residue. Depending on the minerals and the type of equipment, pyrometallurgical recovery may include roasting, smelting, converting, and fire refining. The fire-refined copper is electrorefined using electric current. The impure copper gradually dissolves from the anode and deposits onto the cathode as 99.99% pure copper. The cathode copper is then melted and cast. Hydrometallurgical processes are applied mainly to oxide ores and low-grade sulfides. Copper is recovered from the pregnant leaching solution through precipitation or solvent extraction. Copper is deposited from the loaded solution produced by solvent extraction using electrowinning.

Electrorefining is the final step in the smelting process for the recovery of copper from sulfide ores. The first commercial copper electrorefinery was built in Pembrey, Wales, in 1869, on the basis of patents issued to Elkington in 1867. The Pembrey plant was producing copper of the order of 6000 t year^{-1} by about 1903 and finished operation in the 1910s [18]. This was followed in 1876 by the Norddeutsche Affinerie plant in Hamburg, Germany, which is still operating today.

Copper electrorefining plays a major role in the production and recycling of copper. In the production of copper, copper-bearing sulfide concentrates are first smelted to copper matte. The molten matte is oxidized to blister copper by a Peirce-Smith converter and the blister copper is fire refined and cast to copper anodes. Blister copper contains about 99% copper and impurities such as arsenic, bismuth, iron, nickel, lead, antimony, selenium, tellurium, and precious metals. It is cast into flat anodes, most often on a rotating horizontal wheel. The mold shape includes lugs by which the anodes are supported in the cells. The anode thickness is controlled so that each anode has the same lifetime in the electrorefining cell.

A growing percentage of the world's copper is produced by the leach, solvent extraction, and electrowinning route. The idea of capturing leach solutions rich in copper and passing them directly to an electrowinning plant dates back to the turn of the twentieth century where the first commercial application took place in 1915 at Ajo, Arizona, and La Exotica (Chuquicamata), Chile. The copper produced was of inferior quality, and it had to be reprocessed through a smelter and refinery. As a result, hydrometallurgical processing did not advance until solvent extraction was developed. By using solvent extraction, it was possible to produce a clean, concentrated solution for the electrowinning process. Solvent extraction was first developed during World War II to separate and concentrate uranium, but it was not until the 1960s that selective chemicals were developed for copper [19].

Solvent extraction and electrowinning (SX-EW) is a hydrometallurgical process involving three fundamental steps: ore leaching, solvent extraction, and electrowinning. The first step consists of leaching the copper ore with a weak acidic solution, which, typically, is sulfuric acid. The solution is recovered and mixed in the solvent extraction stage with an organic solvent to selectively remove copper from the acidic leach solution. The organic solvent is then separated and the copper stripped from it with a fresh acidic solution to produce a concentrated, relatively pure copper electrolyte for the final electrowinning stage. In electrowinning, the copper-rich solution is filtered to remove organics, heated, and passed through electrowinning cells, where the copper precipitates

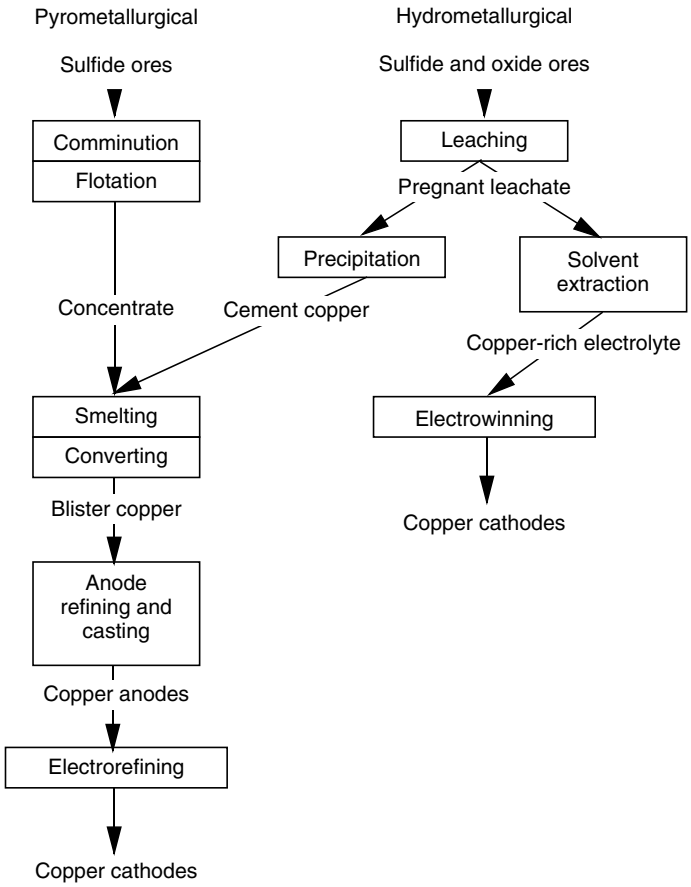


Fig. 14 Copper production technologies.

on copper or stainless steel starter sheets to form cathodes, Fig.14.

4.1.3.1 Copper Electrorefining

The major technical factors in electrorefining are cathode purity, production rate, and specific energy consumption. These factors are influenced primarily by anode quality, electrolyte conditions, and cathode current density. The electrolysis is performed in a solution of copper sulfate and sulfuric acid with a nominal composition of 40–45 g L^{-1} copper and 160–200 g L^{-1} sulfuric acid at 60–66 °C with a current density of 180–300 A m^{-2}. The life of the anode is from 8 to 28 days. Approximately 2 to 10 kg of anode slime per ton of anode remain [20]. The cathodes contain 99.99% or more copper with about 20 ppm or less of metallic impurities. In copper electrorefining, the cell voltage is usually 0.28–0.30 V. The cathodic current efficiency is 92–98%.

Electrorefining consists of placing impure anodes and thin stainless steel or copper cathodes in electrolyte and passing direct electric current between them. The electric current causes copper to dissolve

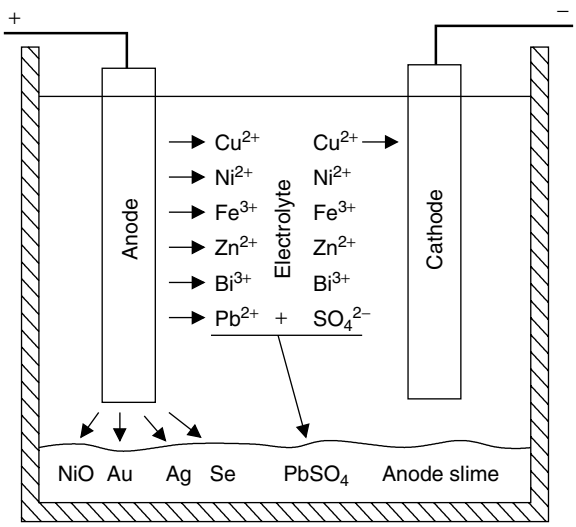


Fig. 15 Principle of copper electrorefining.

from the anodes into the electrolyte and to deposit from the electrolyte on to the cathodes. Some of the impurities do not dissolve (Au, Pt metals, Pb, S, Se, Te, and most Ag). Those impurities that dissolve (e.g. As, Bi, Co, Fe, Ni, and Sb) are continuously removed from the electrolyte. Organic additives are added to the electrolyte to ensure that the deposited copper is dense and smooth (Fig. 15).

The behavior of impure metals depends on their equilibrium potential compared with that of copper:

- Ag, Au, and Pt are more noble than copper and, therefore, will not dissolve anodically. They will remain as metals in the anode slime.
- As, Sn, Bi, and Sb dissolve anodically but will precipitate in the electrolyte as oxide or hydroxide compounds and form part of the anode slime. Because the standard electrode potentials of Bi, Sb, and As are very close to that of copper, they may deposit on the Cu cathode and affect cathode quality.
- Pb dissolves anodically but will form insoluble lead sulfate in the electrolyte. The sulfate will fall to the slime at the bottom of the cell.
- Fe, Ni, Co, and Zn all dissolve anodically and remain soluble in the electrolyte. These elements will reduce at potentials more negative than that at which copper deposits and, therefore, they remain in the electrolyte.
- Nonmetallic impurities, such as NiO and "Kupferglimmer" (Ni–Cu–Sb oxide), do not dissolve and remain in the anode slime.

With or without additives in the electrolyte, the charge-transfer Cu/Cu^{2+} occurs in two consecutive one-electron steps, with Cu^+ being formed as an intermediate (Eq. (20)). The reduction of cupric ion is a slow reaction, whereas the reduction of cuprous ion to metallic copper is a fast one. This means that without additives the formation of Cu(I) ions is the rate-determining step.

$$Cu^{2+} \longrightarrow Cu^+ \quad \text{slow}$$

$$Cu^+ \longrightarrow Cu(ads) \longrightarrow Cu \quad \text{fast} \quad (20)$$

The equilibrium between the Cu(II) and Cu(I) ions and metallic copper is also important. It is described by the disproportionation reaction (21)

$$Cu^{+} + Cu^{+} \longleftrightarrow Cu^{2+} + Cu \qquad (21)$$

Copper can be oxidized by air after reaction (22)

$$2Cu^{+} + 2H^{+} + 0.5O_2 \longrightarrow 2Cu^{2+} + H_2O \qquad (22)$$

The reactions (20) to (22) form the copper equilibrium on the electrode surfaces. Concentration of Cu(I) on the cathode surface affects the deposition rate. The maximum net rate of Cu^{+} production is at about −50 mV versus $Cu/CuSO_4$ and at higher overpotentials it decreases. Disturbing the Cu(II)—Cu(I)—Cu equilibrium can cause the formation of copper powder, but this is more a problem on the anode. For the current densities commonly used in electrorefining, the cathode overpotential is between 50 and 100 mV. The system is mainly charge transfer controlled and the effect of mass-transfer polarization is small. If Cu(I) concentration on the cathode surface decreases, mass-transfer polarization will increase, causing more uneven deposit.

In order to produce a cast anode that is dimensionally consistent and smooth surfaced, some oxygen must be present in the molten copper. During cooling, this oxygen forms cuprous oxide (Cu_2O). During the electrorefining process, Cu_2O dissolves (23)

$$Cu_2O + H_2SO_4 = CuSO_4 + H_2O + Cu \qquad (23)$$

Because of this chemical dissolution, copper builds up in the electrolyte. In order to maintain a constant copper concentration in the electrolyte, continuous electrolyte bleed is decopperized with insoluble anodes to maintain the electrolyte copper level.

4.1.3.1.1 **Technology of Copper Refining**

Figure 16 shows a general flowsheet of the copper-refining operation. The electrorefining process itself is quite simple, but to maintain that process, many other unit processes are needed. The anodes come from casting. The undersides of the lugs are machined to obtain a smooth surface for good connection and vertical position. The weight of the anodes is between 300 and 375 kg. The anodes are loaded to the electrorefining cells, and if the plant uses copper starter sheets, some of the anodes are used for that purpose. The anodes are refined to cathodes. The anode scrap is returned to anode casting and cathodes go for further processing. The electrolyte is constantly purified and anode slimes are collected for the recovery of valuable metals. For every ton of anode copper, 800–870 kg of cathode copper is produced, the anode scrap makes up 130–180 kg, 5–15 kg of copper is enriched in the electrolyte, and impurities in the anode slime make up 2–10 kg.

Typical anode and cathode analyses are shown in Table 5. Comparisons are made to the ASTM standard B 115-91, grade 1 cathode copper and European Standard EN1978:1998, Cu-CATH-1, which is the same as the London metal exchange (LME) grade A copper.

The electrorefining cells are usually 3–6 m long, 1.0–1.2 m wide, and 1.1–1.4 m deep. The electrolyte flows from one end of the cell to another. The temperature is usually maintained at 60–65 °C at the cell inlet, falling to 55–60 °C at the cell outlet. The circulation rates are typically 4–6 L min^{-1} for each cubic meter of cell volume, a cell achieving a nominal

change of electrolyte every few hours. This electrolyte flow helps control temperature, prevents compositional gradients, replenishes additives, and serves as a bleeding-off stream for electrolyte purification. The cells have 30–46 anodes and the number of cathodes is one more or one less. There is also an additional anode or cathode at the end of the cell. The electrode spacing is 90–100 mm between anode centerlines. Longer cells with about 80 electrode pairs and wider cells (Jumbo cells) are also used. Longer cells, for example the Outokumpu system, require closer control of electrolyte flow and better current delivery system to the electrodes. The Outokumpu system uses polymer concrete cells with 80 permanent cathodes at 100-mm spacing [22]. Jumbo cells are very wide cells (up to 28 m) with several electrode racks, each rack having 40 electrode pairs. The racks are connected in the usual Walker current feed design. The electrolyte flow is from one side of the cell to the other [23].

The anodes remain in the electrolytic cells until they are almost completely dissolved, usually for two batches of cathodes. They are then replaced with new ones. Typically, 13–18% of the

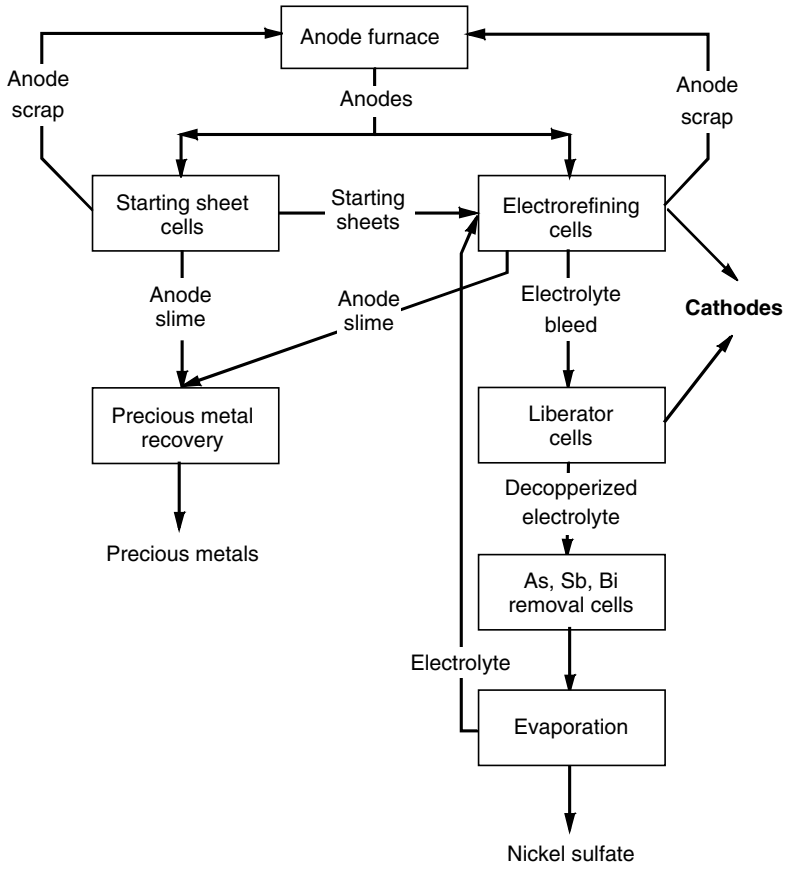


Fig. 16 Typical flow sheet for a copper refinery.

Tab. 5 Typical anode and cathode copper analyses [21]

Element	*Anode concentration variation* [$\mu g\ g^{-1}$]	*Attainable concentrations* [$\mu g\ g^{-1}$]	*ASTM B115-91, grade 1, max.* [$\mu g\ g^{-1}$]	*EN1978:1998, Cu-CATH-1, max.* [$\mu g\ g^{-1}$]
Cu	98.5–99.6%	Balance	Balance	Balance
Se	10–3000	<0.2	2	2
Te	–	<0.3	2	2
Bi	–	0.03–0.3	4	2
Total Se+Te+Bi				3
Cr	–	–	–	–
Mn	–	–	–	–
Sb	100–1500	0.5–1.0	4	–
Cd	–	–	–	–
As	100–4000	0.2–0.7	5	–
P	–	–	–	–
Total Cr+Mn+Sb+Cd+As+P				15
Pb	100–4000	0.1–0.3	5	5
S	–	3.0–6.0	15	15
Sn	10–4000	<0.2	5	–
Ni	100–5000	<0.3–0.5	10	–
Fe	–	0.5–1.0	10	10
Si	–	–	–	–
Zn	–	–	–	–
Co	–	–	–	–
Total Sn+Ni+Fe+Si+Zn+Co				20
Ag	100–3000	4.0–15	25	25
O	1000–4000	–	–	–

anode remains as scrap. The Jumbo cells use lighter continuously cast anodes to minimize accidental damages to tank lining. The anodes are used for one batch of cathodes and the anode scrap percentage is 30% [24]. The undissolved anode scrap is washed and recast as new anodes.

The cathodes are grown for 5–14 days. The cathodes weigh 40–160 kg [20]. After being removed from the cell, the cathodes are thoroughly washed to remove the electrolyte that would otherwise contaminate the copper during melting. Two types of cathodes are commonly used. Cathode copper can be deposited on consumable thin copper starter sheets or on reusable stainless steel or titanium cathodes. Currently, two-thirds of world copper refineries use starter plates and one-third use permanent cathodes [20].

The copper starter sheet is the original cathode construction. The starter sheets are deposited on copper, stainless steel, or titanium cathodes for one to two days and then mechanically stripped. The 1 mm thick starter sheets are cut to operating cathode size, straightened, and fitted with loops for hanger bars. After the production cycle, the removed cathodes are replaced by new starter sheets. A critical aspect of the cell operation is a starting sheet with good quality, straight, square edges, and desired surface roughness.

The major advantages of the reusable cathode blanks are the avoidance of starter sheet manufacturing and improved verticality. Improved cathode verticality improves primary current distribution in the cell, and this has led to fewer short circuits and allowed greater current densities. In the case of reusable cathodes, the copper deposits are mechanically removed and the plates are returned to the cells. There are two major permanent cathode technologies, the ISA process and the Kidd process. The ISA cathode consists of a stainless steel hanger bar welded to a stainless steel plate. The bar and part of the plate are coated with copper to ensure electrical conductivity and corrosion resistance. Traditionally, the ISA process uses wax to prevent copper deposition at the bottom and produces one cathode sheet on each side of the blank. A new system, ISA 2000, has eliminated the use of wax. The Kidd cathode is made by welding a copper hanger bar and stainless steel plate. The Kidd process leaves the bottom exposed, which creates two sheets connected at the bottom. In both the processes, durable plastic edge strips mask the sides of the plate.

During operation, the electrolyte becomes more concentrated with copper, nickel, and arsenic and depleted with sulfuric acid. Floating slimes are composed of antimony arsenate, $SbAsO_4$, and bismuth arsenate, $BiAsO_4$. These can float to the cathode, causing contamination. Soluble impurities are removed from the electrolyte by continuously bleeding a portion of the electrolyte through a purification circuit. The impurity level in the anode determines the volume of electrolyte that must be removed for impurity control. Usually, the elements that control this bleed volume are either arsenic or nickel. The bleed volume is based on the quantity of the controlling impurity and the impurity concentration that is tolerable in the electrolyte. The purification of the bleed stream is done in three sequential steps. The excess copper is first removed and recovered, usually by electrowinning in ''liberator'' cells. Then As, Sb, and Bi are removed via electrowinning to an impure Cu deposit. As the copper concentration decreases, the deposition potential gradually increases until first bismuth and then antimony are reduced to the elemental form at the cathode. Because arsenic deposits at or near the hydrogen potential, the electrical energy requirements are high, and the formation of highly toxic hydrogen arsenide (arsine, AsH_3) may occur. After the removal of arsenic, the nickel sulfate–sulfuric acid solution passes to nickel recovery and acid treatment. Nickel is conventionally recovered by the evaporation of the electrolyte to concentrate and crystallize the nickel sulfate [25].

The insoluble anode impurities are collected as anode slimes from the bottom of the cells and from washing the scrap anodes. The slimes are treated in a precious metals plant for recovery of gold, silver, and other valuable by-products. In conventional slime treatment, the first step is usually the decopperizing in an autoclave at high temperatures and pressures in sulfuric acid media. The filtrate contains tellurium and chlorides. Tellurium is removed by copper, after which the filtrate returns to electrolysis. The slimes are smelted to Doré metal and cast to anodes. The anodes are processed by silver electrolysis. The slime of silver electrolysis contains gold and PGMs. The slime is collected and leached with hydrochloric acid and a strong oxidant. The dissolved gold is precipitated with SO_2 gas, and the PGMs can be cemented.

Tab. 6 Typical voltage components in a copper electrorefining cell [5, p. 603]

Potential source	*Overpotential range [V]*
Electrolyte resistance	0.11–0.13
Cathode overpotential	0.04–0.08
Anode overpotential	0.03–0.06
Contact resistance	0.01–0.02
Total	0.19–0.29

4.1.3.1.2 **Energy Consumption** In electrorefining, the thermodynamic cell voltage is zero. An applied voltage is only required to polarize the electrodes and overcome ohmic resistances. The typical voltage components in the copper electrorefining cell are shown in Table 6.

Specific electrolytic energy consumption is relatively low, around 0.21–0.24 kWh kg^{-1} cathode copper. The electrical energy consumption including losses in the hardware is 0.26–0.28 kWh kg^{-1}. In modern refineries, the cathode current efficiency can be as high as 98–98.5%, the anode current efficiency being slightly higher [5, p. 603]. The anode current efficiency is higher because of the chemical dissolution of cuprous oxide (Cu_2O).

The main component in cell voltage is electrolyte resistance. This depends on the electrolyte resistivity and interelectrode gap. The anode spacing from center to center in world refineries varies between 85 and 115 mm [20]. The interelectrode gap must not be too small as anode slime begins to collect on the cathode. The electrolyte resistivity is a complex function of the electrolyte composition and temperature. Generally, increasing temperature and sulfuric acid concentration lower the resistivity, whereas increasing metal ion concentration increases the resistivity (24) [5, p. 601].

$$\rho = 1.1 \cdot [1.0 - 0.007 \cdot (T - 55) - 0.001 \cdot (H_2SO_4 - 200) + 0.0066 \cdot Cu + 0.0076 \cdot Ni + 0.0082 \cdot Fe + 0.0007 \cdot As \quad (24)$$

where ρ is resistivity in ohm centimeter, T is temperature in degrees centigrade, and concentrations are in grams per liter.

Lowering the electrolyte resistance by varying its composition does not offer many possibilities to work with. The concentration of copper and temperature must be in the usual range to get a good quality deposit. Increasing the concentration of sulfuric acid will increase the chemical dissolution of the anode. The most important method to improve electrolyte conductivity is often by limiting the nickel concentration.

4.1.3.1.3 **Additives** Additives, such as chloride, thiourea, glue (gelatin) and Avitone, are used in electrorefining as grain refiners and leveling agents. Glue acts as a leveling agent. Thiourea, Avitone (trade name for a group of chemical compounds based on hydrocarbon sodium sulfonates), and chloride act as grain-refining agents. Although the exact mechanism and synergism of the leveling action of additives is not well known, their effect on cathode quality and the influence of the electrolysis parameters on the leveling action are documented [26]. The additives allow the production of smooth dense cathodes since morphology is extremely important in maintaining cathode purity. Nodules and dendrites are the most common surface defects that lead to impurities in the cathode. The glue dosage ranges from 30–40 to 190–200 g t^{-1} of copper. The

thiourea dosage ranges from less than 10 to more than 100 g t^{-1} of copper.

The glue adsorbs mainly at locations with high electric field strength, such as edges, nodules, and needles. Adsorption of the polar glue molecules on these areas leads to the formation of an isolating layer, and the field strength decreases. As a result, growth of the protrusion stops. After the copper layer has grown up to the level of the protrusion, the field strength equalizes, the glue is desorbed, and it can adsorb again in another area with high field strength. The glue acts as an inhibitor, which leads to the deposition of the base-oriented structure. Low concentrations of thiourea depolarize the cathode and the deposit structure changes to a field-oriented structure. Thiourea can also react with glue, which results in the reduced cathode-polarizing effect of glue. The deposit structure is still compact and the growth proceeds with the base-oriented structure with Cu crystals having clearly restricted length. Chloride ion is added to modify the morphology, structure, and mechanical properties of the deposits. The chloride ions give a smooth deposit surface structure (Fig. 17) [27].

The correct addition of leveling agents is critical to the production of smooth and dense cathodes. The higher the current density, the more critical the additive levels become. In the case of glue, the cathodic polarization values are indicative of adequate active glue levels. As the glue loses it activity, the cathodic

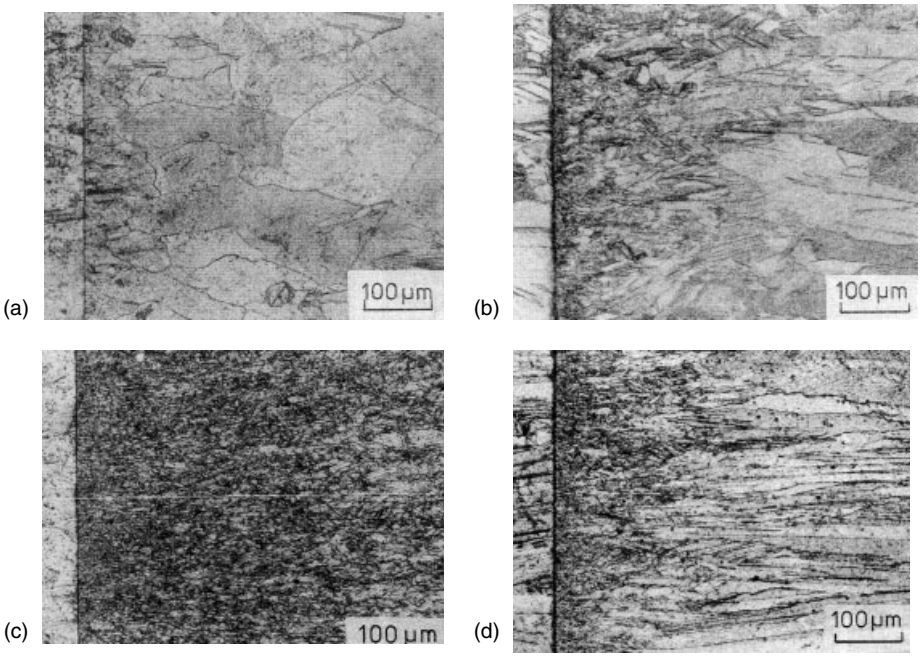


Fig. 17 Copper deposit morphologies from crosscut samples. Electrolyte contains 180 g L^{-1} H_2SO_4, 45 g L^{-1} Cu^{2+}, and 10 g L^{-1} Ni^{2+}. $T = 60\,^{\circ}C$ and $i = 350$ A m^{-2}. (a) Without additives, (b) with 2 mg L^{-1} glue, (c) with 2 mg L^{-1} thiourea, and (d) with 2 mg L^{-1} glue, 2 mg L^{-1} thiourea and 30 mg L^{-1} Cl^{-}.

polarization decreases. Today, the CollaMat method is widely used to continuously measure the active glue concentration in the electrolyte. The CollaMat system is a proprietary real-time glue-monitoring apparatus. With the CollaMat process, copper electrolysis is performed under certain conditions in a specially constructed measuring cell through which the electrolyte flows continuously during the measuring cycle. The glue content is calculated from the kinetic data compiled during the measuring process [26]. The Reatrol Process is a patented technique, which utilizes differential pulse polarography to ascertain the active concentration of thiourea by complexation with a dropping mercury electrode [28].

In order to achieve high production rates, high current densities are desirable. High current densities can cause two major problems. Increasing the current density causes higher impurity levels in the cathode as increased roughness promotes the occlusion of floating slime and electrolyte. The anode can passivate when impurities gather on the anode surface, causing the crystallization of copper sulfate. Many copper producers have increased the current density to 300–350 A m^{-2}. Copper electrorefining under conventional electrolysis conditions is possible up to 1000 A m^{-2}, in principle. However, increasing the current density will result in large dendrite formation on the cathode and anode passivation. In the current density range 700 to 1000 A m^{-2}, periodic oscillations of the cathodic polarization and bath voltage reveal instabilities in the traditional additive system of glue, thiourea, and chloride. A significant increase of current density in copper electrorefining requires alternative additives [29].

4.1.3.2 Sulfate-based Copper Electrowinning

About 20% of primary copper is obtained by leaching the oxide and supergene sulfide ores [17]. On a capacity basis, Chile is the largest producer of electrowon copper [30]. Sulfate processes using oxygen as the oxidant are the most common methods for copper leaching. It is generally necessary to use pressure leaching at high temperatures (150–225 °C) to achieve rapid leaching rates of 1–2 hours. The atmospheric leaching processes with fine grinding to less than 20 μm require 12–24 hours for high Cu extraction and biological leaching processes typically require several days [31]. In dump or heap leaching, the pregnant leach solution contains 0.5–5 g L^{-1} Cu; in bacterial leaching, 20–50 g L^{-1} Cu; and in pressure acid leaching, 20–90 g L^{-1} Cu.

The most important copper electrowinning production method is heap leaching of oxide copper ores with recovery by solvent extraction and electrowinning (SX-EW process). The general flowsheet of an SX-EW process is shown in Fig. 18. It is a low-cost method of copper recovery. This technology has recently been applied successfully to mixed oxide and chalcocite ores, notably in Chile. Currently, there are significant development efforts underway to try to extend heap leaching to chalcopyrite ores.

The copper electrowinning process requires concentrated solutions to improve mass transport and increase the solution conductivity. The pregnant leach solutions from leaching are too dilute and too impure for the direct production of high-purity cathodes. Electrowinning from these solutions would give impure, dendritic deposits. Solvent extraction provides the means for producing pure,

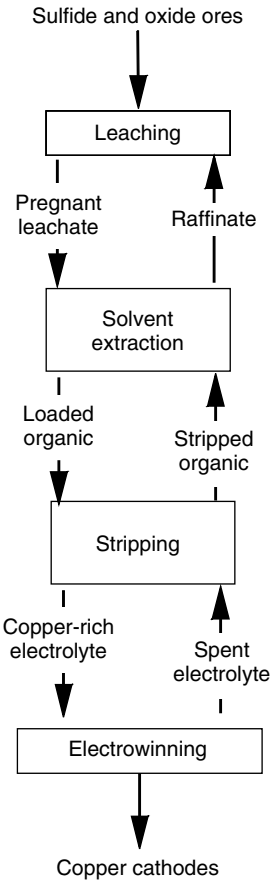


Fig. 18 General flowsheet of copper leaching, solvent extraction, and electrowinning process.

high-copper electrolytes. The copper is deposited from a solution usually containing 25 to 60 g L^{-1} copper as copper sulfate and 50 to 180 g L^{-1} sulfuric acid, at a temperature of 50–60 °C. The same types of additives as in copper electrorefining are used. Addition of cobalt to the electrolyte is used to lower the anode overpotential. The typical current density is 300 A m^{-2}. The cell voltage is about 2.0 V and the energy consumption is about 2.0 kWh kg^{-1} [30].

Over 70% of the electrowinning plants use polymer concrete cells. The cell length can vary from 3 to 17 m. The cathode center-to-center distance is 90–110 mm. In copper electrowinning, the oxygen-evolving anodes are usually made from lead. The most common anode type is the rolled lead–calcium–tin anode. The traditional cathodes for copper electrowinning were thin copper starter sheets made by plating copper onto titanium or stainless steel electrodes. Two-thirds of the copper electrowinning plants use stainless steel permanent cathode technology. The cathode copper is stripped from the stainless steel in automated machines [30].

The Sherritt–Cominco process was developed as an alternative to smelting of sulfide concentrates. The key feature of the Sherritt–Cominco process is the removal of iron before the leaching of copper. The concentrate is thermally activated and leached with sulfuric acid to dissolve iron. Iron is precipitated as jarosite. The leach residue is pressure leached with oxygen in acid solution to convert copper sulfides to copper sulfates. The solution is purified by the coprecipitation of Te, As, Bi, Sb, Pb, and Se with Fe_2O_3. The electrolyte contains 30 g L^{-1} Cu and 140 g L^{-1} sulfuric acid. The electrowinning is done in covered air-sparged cells using lead anodes and titanium cathodes. Current density in electrowinning is high, 430–650 A m^{-2}. The cell voltage is 2.5 V and current efficiency over 90% [32].

4.1.3.3 Halide-based Copper Electrowinning

Chloride leaching of copper sulfides has received a great deal of attention in past decades. The chloride processes have distinct advantages over sulfate-based processes. The chloride systems have faster kinetics, making atmospheric leaching

more practical. Almost all the sulfides are oxidized to elemental sulfur. The higher solubility of metals means smaller solution volumes in subsequent steps [33]. The major difficulties are the need for special construction materials due to the corrosivity of chloride solutions leading to relative high capital and maintenance costs for the plant. Most of the chloride-based processes can be grouped into three main classes:

1. Processes that use chloride leaching, solvent extraction to sulfate solution, and electrowinning from sulfate solution.
2. Processes that use ferric chloride or cupric chloride to produce cuprous chloride solution. Copper is recovered by chemical precipitation.
3. Processes that use ferric chloride or cupric chloride to produce cuprous chloride solution. Copper is recovered by electrowinning in diaphragm cells, and the leachant is regenerated at the anodes.

Electrowinning directly from chloride solutions usually does not produce solid cathode copper. The exchange current density of Cu^{+}/Cu equilibrium in pure chloride solution is very high, 240 mA cm^{-2} [34]. According to the theories of Fischer [7] and Winand [8] on electrodeposition, the reaction inhibition is low and so the deposit morphology becomes easily powdered. The electrowinning of copper from chloride solutions often results in impure deposits. In a cuprous chloride solution, the discharged copper ions are monovalent, so the electrochemical equivalent is half of that of divalent copper in sulfate solutions. The anodic overpotential of chlorine evolution is 0.5–0.6 V lower than that of oxygen overpotential with lead anodes in sulfate solutions. The lower electrochemical equivalent of cuprous ion and lower anode overvoltage would significantly decrease the energy consumption. However, the current efficiency in cuprous chloride electrolytes is low if the electrolyte contains even small amounts of cupric ion. The cupric ion is a powerful oxidant, and it will again leach the cathode deposit.

The Duval Corporation's CLEAR (copper leach, electrolysis, and regeneration) process was operated in a 100 t day^{-1} copper leaching/electrowinning plant in Arizona, USA, between 1976 and 1982 [35]. The concentrates were leached countercurrently in two stages. The first leaching stage used hot cupric chloride at atmospheric pressure to yield a cuprous chloride solution. After clarification, the pregnant leaching solution was sent to copper electrowinning. Copper was deposited as a powder. The copper is electrowon from nearly saturated chloride solution. The electrowinning cell is of unique construction. Metal cathodes are suspended in the solution above a continuous conveyor belt within the tank. Nonmetallic anodes are suspended in diaphragm bags [36]. The copper that was produced in CLEAR process by electrowinning contained Ag and other impurities that necessitated its treatment as blister copper, that is, it had to be melted, cast into anodes, and refined electrolytically [35]. The commercial operation had several major problems including gypsum fouling in the tankhouse, silver contamination of the final product, and washing of the final product.

In the Cuprex process, the copper sulfide concentrates were leached with a sodium chloride and ferric chloride solution in two stages to produce a cupric chloride solution. Copper is extracted by solvent extraction and recovered as powder by electrowinning. The process used

Metclor cells with DSA-type anodes, hollow titanium cathodes, and cation-selective membranes. Copper is deposited on the cathode and chlorine is evolved on the anode. Sodium ions transport through the membrane to the catholyte. The operating principle of the Metclor cell was quite close to that of the chlor-alkali diaphragm cell. Copper was deposited at the current density of 1500 A m^{-2}. The spent catholyte contained cupric, cuprous, and sodium ions. Chlorine generated at the anode was recovered and used to reoxidize the cuprous chloride generated in the catholyte during electrowinning. The recovered anolyte was depleted from copper and used again as anolyte [37].

The Intec Copper Process is a chloride-based method of converting copper sulfides to high-purity copper and elemental sulfur. The main components of the leaching solution are sodium chloride and bromide. The lixiviant, $BrCl_2^-$, also known as *Halex*™, is produced anodically in the electrowinning cell. The leaching is carried out at atmospheric pressure using air as the oxidant. The process utilizes a three-stage countercurrent leaching at 80–85 °C. Copper is leached in stages 1 and 2 and gold is leached in stage 3. The pregnant solution is purified in several stages. First, all the copper is converted to cuprous by contact with copper powder. Second, silver is removed by an amalgam process using mercury and aluminum additions. The final step is neutralization with limestone to pH 4.0–4.5 to precipitate the residual impurities. The purified cuprous chloride–sodium halide solution contains 80 g L^{-1} Cu^+, 250 g L^{-1} NaCl, 30 g L^{-1} CaCl, and 30 g L^{-1} NaBr at 75 °C. The solution is fed to a diaphragm cell where Cu dendrites are deposited on corrugated titanium cathode. The depleted catholyte permeates through the membrane, where the lixiviant is formed and the remaining copper is oxidized to Cu^{2+}. At current densities of 1000 A m^{-2}, the cell voltage is 3.2 V. The current efficiency is 98%; the power consumption is approximately 1.4 kWh kg^{-1}. The dendritic copper produced from electrowinning is collected by a belt conveyor, washed, and dried under inert atmosphere [38].

4.1.4 Nickel Production

World nickel metal production in 2002 was 678 000 tons [39]. Hydrometallurgy has typically been applied to the treatment of nickel–copper mattes, anode nickel, and reduced laterite ore. The sulfide concentrates are usually oxidized by roasting and then smelted to copper–iron–nickel sulfide matte (75–80% Cu–Ni), which is refined or used directly to make Monel metal. Cathode nickel can be produced from a variety of electrolytes, including chloride, sulfate, or a mixed chloride–sulfate. The former two are acid systems used in leaching and electrowinning. Mixed chloride–sulfate electrolytes are used for electrorefining the nickel sulfide matte from the traditional matte-smelting operations.

The first electrolytic nickel refinery plant, treating nickel metal anodes, was built by Hybinette in Kristiansand, Norway, in 1910, and this plant was acquired by Falconbridge in 1928. INCO developed electrorefining of nickel sulfide matte anodes in the 1950s. In a typical electrowinning process, the raw material is first smelted to a matte and then leached in a sulfate or chloride solution. The sulfur of the raw material is oxidized to insoluble elemental sulfur or soluble sulfate. The nickel contents of the mattes treated hydrometallurgically are in the

range 39–79%, while the copper content varies from 1 to 52% and sulfur varies from 4 to 23%. Mattes with less than 7% copper can be leached by ammoniacal pressure leaching, ferric chloride leaching, or direct electrorefining of the matte. High-copper mattes containing over 25% Cu can be dissolved under atmospheric pressure or using autoclaves, by sulfuric or hydrochloric acid. Nickel is recovered from the purified solution either electrolytically as a nickel cathode or by chemical reduction with hydrogen as nickel powder.

4.1.4.1 Nickel Electrorefining

The first nickel electrorefining process was the Hybinette process developed by the Swedish metallurgist Noah Victor Hybinette in 1905 for refining crude nickel anodes. The process forms the basis of modern nickel electrorefining practices. The anodes of the Hybinette process were usually nickel–copper alloys with 84% Ni and 15% Cu. The cathode plates were usually nickel sheets. The electrolysis was done in weakly acidified nickel sulfate and boric acid electrolyte with 50–60 g L^{-1} Ni and 10–20 g L^{-1} H_3BO_3. The Hybinette cell had separated anode and cathode departments, Figure 19. The electrolyte from the anode compartment was purified by the removal of iron and copper. The purified electrolyte was pumped to the cathode compartment which had solution level higher than in the anode compartment. The depleted electrolyte flowed through cloth separators to the anode compartment. At the anode, nickel, copper, and iron dissolved, while the precious metals fell to the anode slime (Fig. 19). The temperature of refining was 55–60 °C. The current density was 110–135 A m^{-2}, cell voltage was 2.3–2.7 V, and the current efficiency was 80–85%. The energy consumption was 3–4 kWh kg^{-1} nickel [40, 41].

4.1.4.1.1 Electrorefining Nickel Metal Anodes

In the refining of nickel metal anodes, the principal reaction at the anode is the dissolution of nickel metal as nickel ions. The principal cathodic reaction is the reduction of nickel ions from solution. Nickel anodes are made by reducing nickel oxide with coke at temperatures up to 1550 °C and casting in molds. The practice is designed to obtain anodes with the desired strength and crystal size.

The anodes contain nickel both as metal and as sulfide. During nickel metal anode refining, the nickel phase dissolves selectively, leaving sulfides undissolved. The increase in sulfur content decreases the anode overvoltage but increases the quantity of anode slime [42].

Depending on the practice, the tank holds 30–40 anodes with an additional anode at the end of the tank to prevent warping of the cathode. The cathode compartments of the cell are segregated from the anode by a porous diaphragm enclosing the cathode. If the soluble impurities reach the cathode surface, they will codeposit with the nickel. To avoid contamination of the refined nickel cathodes, the impure anolyte solution is taken through a purification process. The purified solution is fed into the cathode compartment at a rate sufficient to maintain the rate of nickel deposition. Owing to the diaphragm cloth, the solution level in the cathode compartment is higher than that in the rest of the tank. This hydrostatic force causes the solution to flow from the cathode compartment through the diaphragm into the anolyte. This flow prevents the ingress of anolyte or impurity ions into the cathode compartment. The nickel starting sheets are made by depositing nickel onto permanent stainless steel or titanium cathode blanks, from which the thin sheet is stripped after 24 or 48 hours. The starting

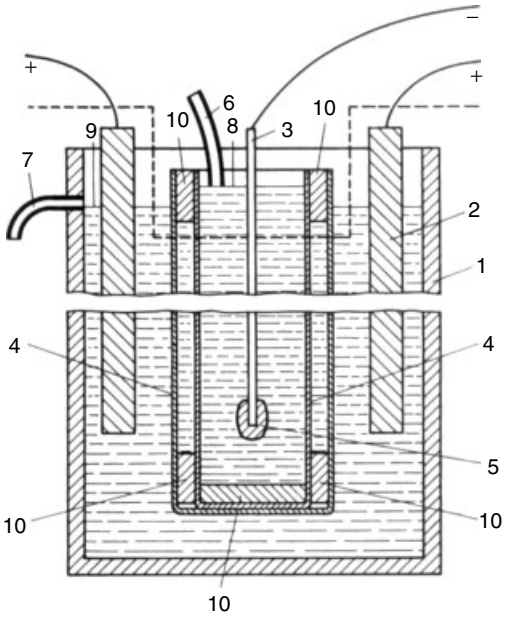


Fig. 19 Hybinette cell for the electrorefining of nickel. 1 – the cell, 2 – anode, 3 – cathode, 4 – diaphragm, 5 – cathode-insulating edge strip, 6 – catholyte feed, 7 – anolyte feed, 8 – catholyte surface level, 9 – anolyte surface level, 10 – diaphragm support [41].

sheets are fitted with suspension loops and placed in the production cells.

INCO produced electrolytic nickel at its refinery in Port Colborne, Ontario, Canada. The production started in 1926. The anodes were made by reducing nickel oxide with coke, and the anodes contained about 93.5% Ni, 4% Cu, and 1% Co. The sulfur content was low, about 0.6%. The approximate composition of the purified electrolyte was 60 g L^{-1} Ni^{2+}, 95 g L^{-1} SO_4^{2-}, 35 g L^{-1} Na^+, 55 g L^{-1} Cl^-, and 16 g L^{-1} boric acid, and the temperature was 60 °C. The current density of the process was 16 A/sq.ft (approximately 170 A m^{-2}) and the cell voltage was about 1.6 V. At the normal cell operating voltage, the principal impurities – iron, cobalt, lead, arsenic, and copper – dissolved into the solution with nickel. Silver, gold, the PGMs, sulfur, selenium, and tellurium fell to the bottom of the cell as an insoluble slime. The produced cathodes had a purity of better than 99.9% Ni [43]. Today, INCO no longer refines nickel in Port Colborne.

Falconbridge Ltd electrorefined nickel anodes in its Kristiansand, Norway plant using the Hybinette electrorefining process until 1981. The raw material was crushed converter matte produced at the Falconbridge, Ontario, Canada smelter. The matte contained 48% Ni, 27% Co, and 22% S. The matte was ground and roasted to oxide. The oxide was leached in sulfuric acid to remove copper and the remaining oxide was melted using coke for reduction. The anode composition was about 76.5% Ni, 17% Cu, 2.0% Fe, 1.4% Co, and 1.4% S. The anode size was about 95 × 60 cm^2, with 5.5 cm thickness and the weight about 260 kg. The cathode starting sheets were slightly larger, 100 × 67 cm^2. The tanks had 39 anodes and 38 cathodes. The operating current density was 180 A m^{-2}. The cathode cycle was typically eight days

and the final cathode weighed 50 kg. The cathode purity was better than 99.95% Ni [43].

The refining of metallic nickel anodes is still practiced in the former Soviet Union. The Severonikel complex at Monchegorsk in the Kola Peninsula treats a blend of high-grade nickel–copper ores, concentrates, mattes, and scrap from the nickel-mining operations at Norilsk and Pechenga. The oxide calcine is reduced with coke to form a metallized oxide, which is melted at 1550 °C in an electric arc furnace for anode casting. The anodes typically weigh 350 kg and contain 86–88% Ni, 2% Co, 4.5% Fe, 4% Cu, and 0.2% S. The electrolyte contains 75–78 g L^{-1} Ni, 130–150 g L^{-1} sulfate, 14 g L^{-1} Cl^-, and 6 g L^{-1} H_3BO_3, at pH 2. The current density is 250 A m^{-2}. The cathodes are enclosed in a synthetic fiber bag supported on a polypropylene frame. The anode cycle is 30 days, while the cathode cycle is 4 or 6 days. The anode slimes are recovered by washing the adhering material off the corroded anodes, and from the bottom of the electrolytic cells, at the end of each anode cycle, for subsequent processing to recover precious metal values [44].

4.1.4.1.2 **Electrorefining Sulfide Matte**

The International Nickel Company developed a method to refine impure nickel sulfide anodes directly to metal, using mixed sulfate–chloride electrolyte [45]. Nickel sulfide (α-Ni_3S_2) anodes can be cast directly from low-copper converter matte or from melted nickel sulfide concentrate produced by the matte separation process. Controlled cooling is necessary to produce anodes with the required mechanical properties. The cooling of anodes can take up to 36 hours. Using nickel sulfide anodes eliminates the intermediate roasting of the sulfide to oxide and reducing the oxide to produce metallic nickel anodes. The anode composition is essentially Ni_3S_2 and nickel alloy with roughly 75% nickel and 25% sulfur. Control of the sulfur content of the anodes is important. If the sulfur content is less than 15%, the metallic nickel dissolves preferentially, leaving undissolved nickel sulfide in the slime. At sulfur contents above 15%, the anode corrodes uniformly with elemental sulfur in the slime. The anodes can contain up to 3% copper as Cu_2S and metallic impurities such as iron, cobalt, lead, zinc, and PGMs.

The principal anodic reactions are the oxidation of the sulfur of the nickel sulfide to elemental sulfur with simultaneous release of metal ions to the solution (25)

$$Ni_3S_2 \longrightarrow 3Ni^{2+}(aq) + 2S(s) + 6e^- \quad (25)$$

The metallic nickel phase dissolves as in the refining of nickel metal anodes. The Ni_3S_2 dissolution by reaction (25) requires 0.2 V higher overpotential than dissolution of the nickel metal. The cathodic reaction is the deposition of nickel. The anode potential is sufficiently high so that some oxygen evolution happens. The unwanted oxygen evolution reaction, together with the electrolytic dissolution of impurities from the anode, lowers the anodic nickel dissolution efficiency so that nickel is deposited from the solution at the cathode faster than it dissolves at the anode. This imbalance is corrected during electrolyte purification by dissolving nickel matte in the electrolyte to neutralize the free acid and replenish the nickel content [43].

Operation of the nickel matte refining cell is similar to that of the metal anode cell except that the anodes are enclosed in bags to collect the voluminous sludge. The cathodes are placed in conventional cathode compartments. The anode spacing

in the matte-refining cell is greater than that in the metal-refining cell, to allow for the increase in volume of the matte anode as it corrodes. The actual cell voltage, including the potential drops across the contacts and through the electrolyte, is about 2.8 V for new matte anodes. As the anodes dissolve, the sludge forms a strongly adherent porous surface layer. As the sludge layer becomes thicker, the electrical resistance between the electrodes increases and the cell voltage increases to about 4.0 V by the end of the anode cycle. The increase in cell resistance also results in increased heating of the electrolyte. At the beginning of the anode cycle, the temperature of the electrolyte leaving the tank is 15 °C higher than that at the time of entering, but at the end of the anode cycle the temperature difference is 25 °C [43].

The INCO, Thompson plant in Manitoba, Canada, electrolyzes 240 kg sulfide anodes in a sulfate–chloride electrolyte. The approximate composition of the electrolyte is 60 g L^{-1} Ni^{2+}, 95 g L^{-1} SO_4^{2-}, 35 g L^{-1} Na^+, 60 g L^{-1} Cl^-, and 16 g L^{-1} H_3BO_4, and the temperature is 60 °C. Nickel, cobalt, and copper dissolve from the anode, while sulfur, selenium, and the noble metals form an insoluble sludge or slime, from which they can be recovered. The anode sludge contains 95% elemental sulfur, sulfide sulfur, nickel, copper, iron, selenium, and precious metals. Nickel is deposited on to pure nickel starting sheets. The anode cycle is 15 days and the cathode cycle is 5 to 10 days. Electrolysis is carried out at a current density of 240 A m^{-2} giving a cell voltage of 3 to 6 V [44, 46].

4.1.4.2 Sulfate-based Nickel Electrowinning

The electrowinning processes use nickel sulfide matte, which is dissolved. To prevent the deposition of impurities, the electrolyte is routed through a purification process before the solution is pumped to the cathode compartment. Nickel electrowinning is done in diaphragm cells using permanent anodes. In the sulfate system, the anodes are usually made of lead but permanent DSA-type anodes can also be used. The anode reaction is oxygen evolution which leads to the formation of hydrogen ions.

In a sulfate electrowinning system, the cathodes are suspended in bags, even though the nickel sulfate solution is purified before it enters the tankhouse. The cathode must be protected from strongly acidic anolyte. A high concentration of hydrogen ions at the cathode would result in hydrogen gas evolution that will reduce the current efficiency and lower the cathode quality.

The energy consumption in electrowinning depends on the thermodynamical voltage, electrode polarization, and electrolyte resistivity. The conductivity of nickel electrowinning electrolytes (catholyte) at 60 °C is typically 120 mS cm^{-1} and for nickel electrorefining electrolytes at 60 °C it is about 200 mS cm^{-1} [47]. The sulfate electrolyte conductivity can be estimated using Eq. (26)

$$\kappa = 114.84 - 0.39359 \cdot [\mathrm{Ni}] + 0.79499 \cdot [\mathrm{H_2SO_4}] + 1.0328 \cdot T \quad (26)$$

where κ is the electrolyte conductivity in millisiemens per centimeter, the concentrations are in grams per liter, and the temperature is in degrees centigrade [48]. Equation (26) indicates that electrorefining and electrowinning at any given sulfuric acid concentration should be carried out at low nickel concentration. Low nickel concentration, on the other hand, would give a low limiting current density.

The OMG Harjavalta nickel refinery in Finland operates the plant developed by Outokumpu and started in 1960. The process uses flash furnace and electric furnace mattes as raw materials. The flash furnace mattes have high iron and low sulfur content and electric furnace mattes have low iron and high sulfur content. The fine ground matte is leached using nickel electrowinning anolyte containing 100 g L^{-1} Ni and 50 g L^{-1} sulfuric acid. In the reactors, the feed is strongly agitated and oxygen is fed to the reactors from below. The temperature in the reactors is 85 °C. Seventy percent of the incoming nickel is leached. Copper is leached and precipitated as sulfides. Cobalt is nowadays separated by solvent extraction. The raffinate from cobalt extraction contains 130 g L^{-1} nickel, sodium sulfate, and boric acid to improve the deposit. The latter additive adjusts the pH to 3.5–3.8. The metal is deposited at 65 °C onto nickel starter sheets and oxygen is evolved at the lead anodes. The sulfuric acid generated at the anode is utilized by recirculating the anolyte to the leach circuits [49–51].

The OMG Harjavalta nickel refinery has 126 electrowinning cells, of which 108 are commercial cells and 18 are starting sheet shells. The older cells have 49 insoluble lead anodes and 48 cathodes and the newer cells have 53 anodes and 52 cathodes each. Figure 20 shows the schematic cell construction used at the Harjavalta nickel refinery. The bagged cathodes are nickel starting sheets made by the deposition of nickel for 48 hours onto a titanium mother blank. The deposits are stripped and automatically made up into starting sheets. The purified solution is fed separately to the cathode compartments at

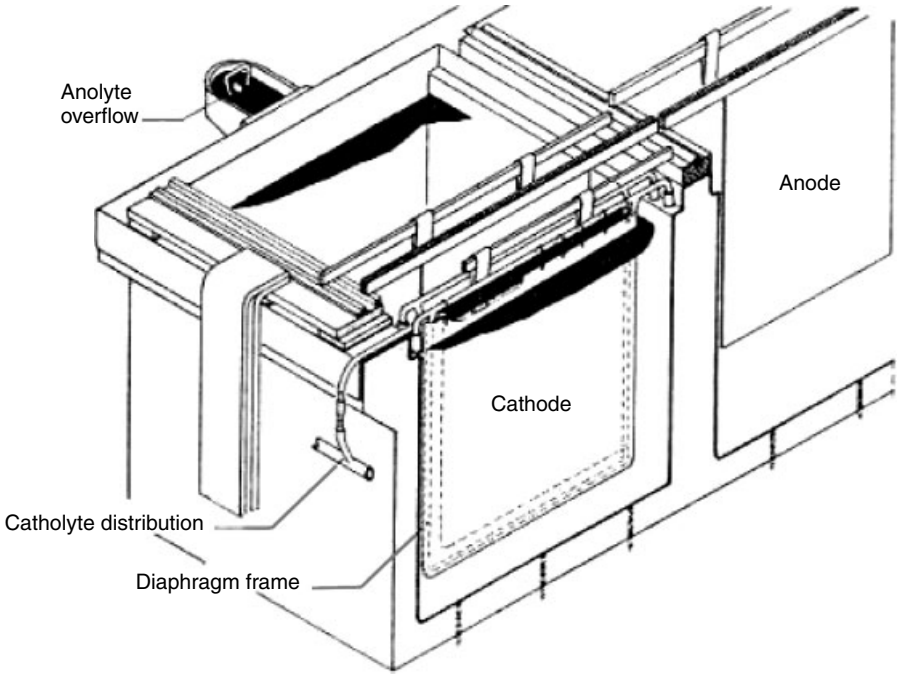


Fig. 20 The cell construction used at the Harjavalta nickel refinery [51].

a rate sufficient to maintain a hydrostatic head of about 30 mm. The current density is 180–200 A m^{-2}, giving a cell voltage of about 3.6 V. The cathode cycle is seven days and the weight of one cathode is 65 kg [51].

In the Sherritt acid pressure leaching process, the high-grade Ni–Cu matte is ground and treated in atmospheric pressure with spent copper electrolyte. Some of the nickel is dissolved in this step. The solution is treated to crystallize nickel sulfate. The leach residue is leached in an acid oxygen pressure leach process ($CuSO_4 - H_2SO_4 - O_2$) at a temperature of 140 °C and a pressure of 550 kPa. In the pressure leach, all the nickel and copper are dissolved and all the sulfur in the matte is oxidized to sulfate. After purification steps to remove selenium, iron, lead, and cobalt, copper is recovered by electrowinning. Zinc from the zinc sulfate is recovered by electrowinning or hydrogen reduction [46].

4.1.4.3 Chloride-based Nickel Electrowinning

The advantages of using chloride electrolytes compared with sulfate electrolytes are higher electrical conductivity, lower electrolyte viscosity, lower overpotential for nickel reduction, and higher solubility and activity of nickel. An important factor is the lower anode potential of chlorine evolution compared with oxygen evolution in sulfate electrolytes using the common lead anodes. Chloride electrolytes require insoluble or dimensionally stable anodes, usually titanium coated with an electroactive noble metal or oxide, and a diaphragm system to collect the Cl_2 gas from the anode. The chlorine liberated at the anode is recycled for use in the leach circuits. In practice, some decomposition of water also occurs to yield small quantities of hydrogen at the cathode and oxygen at the anode.

Falconbridge Nikkelverk A/S in Kristiansand, Norway, uses the chlorine leach and electrowinning process. The matte is first ground and then chlorine leached to dissolve the contained nickel. The chlorine leach solution is oxidized with chlorine gas and neutralized with nickel carbonate to precipitate iron hydroxide. The filtrate is pumped to solvent extraction of cobalt. The cobalt solution is sent to cobalt electrowinning. The raffinate contains 220 g L^{-1} Ni and it is diluted to 85 g L^{-1} Ni for Pb and Mn removal, and after that, it is diluted to 60 g L^{-1} Ni for electrowinning. Typical operating parameters for electrowinning are current density of 220 A m^{-2} and current efficiency of 98–99%. The catholyte contains 60 g L^{-1} Ni and the anolyte contains 54 g L^{-1} Ni and the electrolyte temperature is 60 °C. The refinery produces both regular nickel cathodes and Falconbridge Crowns™. The regular cathodes are produced by plating onto nickel starting sheets. Crowns weighing 30–60 g are round pieces produced by plating nickel onto nonconductive cathode blanks with evenly spaced conductive spots [52, 53].

Nickel electrowinning is done in diaphragm cells using DSA anodes and nickel starting sheets. Each cell contains 46 anodes and 45 cathodes with an anode spacing of 145 mm. In the Falconbridge system, the anodes are enclosed in a polyester diaphragm bag to collect the chlorine gas generated in the anode reaction. Figure 21 shows the Falconbridge nickel electrowinning cell design. The chlorine gas is channeled to a vacuum hood fitted to the top of the anode. Ducts, connecting the top of the hood with a manifold that runs alongside the tank, allow the

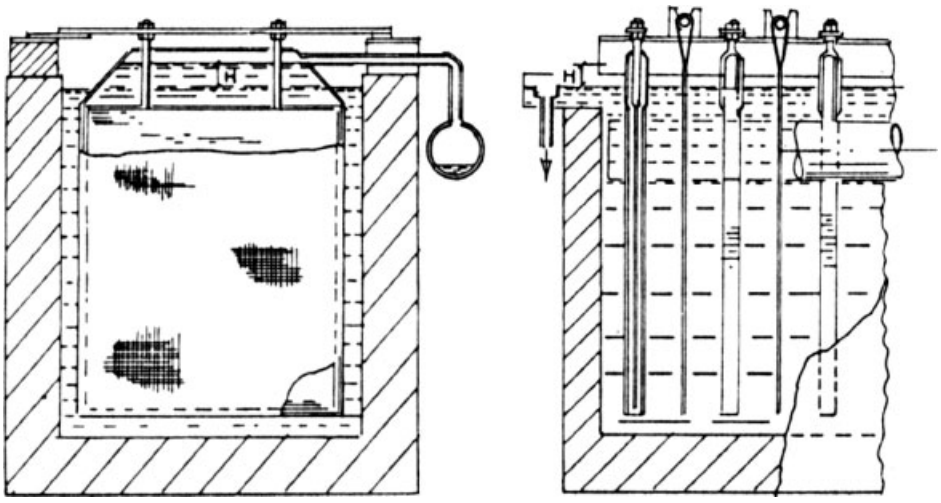


Fig. 21 Falconbridge nickel electrowinning cell design [52].

withdrawal of both chlorine and anolyte by suction. Purified solution is fed in at one end of the cell, and the solution level is held constant by maintaining a small overflow at the other end, while the greater portion of the electrolyte is withdrawn as anolyte with the chlorine gas.

The Sumitomo Metal Mining Co. formerly operated nickel metal and matte electrorefining processes. In 1992, they started a matte chlorine leach and electrowinning (MCLE) process to treat their nickel matte at Niihama in Japan. In the MCLE process, metallic elements are leached with chlorine gas. Cupric ions are used as an oxidant in leaching. Copper ions are removed by electrowinning. Nickel and cobalt are first solvent extracted together. Stripping is done with dilute HCl to produce cobalt and nickel chloride solutions. The Co and Ni are then separated also by solvent extraction. The nickel electrowinning is done in purified nickel solution at pH = 1.5 and temperature of 60 °C. The process uses DSA anodes with diaphragm bags. Chlorine gas and anolyte are sucked out and sent for chlorine recovery. The current density in Ni electrowinning is 230–250 A m^{-2} and cell voltage is 3.0–3.1 V [54].

SLN's Sandouville nickel refinery at Le Havre, France, produces electrowon nickel from matte containing 75% nickel. The matte contains virtually no copper. The matte is first finely ground and then leached with ferric chloride in the presence of chlorine. Nickel, cobalt, and iron are converted into chlorides, while sulfur remains in the elemental state. The solution of nickel, cobalt, and iron chlorides is filtered to remove sulfur and other insoluble materials and then put through processes of extraction and purification. Iron and cobalt are extracted by using selective organic solvents. This yields a solution of pure cobalt chloride and a cobalt-free solution of nickel chloride. The nickel solution is electrowon using insoluble anodes and nickel starter sheets [46].

4.1.4.4 Nickel Electrolyte Purification

For the production of nickel cathodes, it is essential that the solution is purified before

it is pumped to the cathode compartment. Treatment of electrowinning anolyte or spent electrolyte varies, depending on the type of leaching circuit used. The purification unit processes and their order depend on the impurity levels in the anode.

Cobalt is present in almost all nickel ores. Previously there was no use for cobalt and as its value was less than that of nickel, it was not separated. Therefore, old nickel alloys contain much more cobalt than those of the present day. Two chemical methods are based on the oxidation of cobalt to three-valent ions. Cobalt is precipitated as cobalt(III) hydroxide by treating the solution with chlorine to oxidize the cobalt(II) to cobalt(III) and adding nickel carbonate to control the pH to 4 (27)

$$2Co^{2+} + Cl_2 + 3NiCO_3 + 3H_2O \longrightarrow 2Co(OH)_3 + 3Ni^{2+} + 2Cl^- + 3CO_2 \quad (27)$$

Chlorine cannot be used in pure sulfate electrolyte; nickel(III) hydroxide is normally used (28). Other oxidants that can be used to remove cobalt from the sulfate system include peroxide, persulfate, and ozone [5, p. 803]

$$Co^{2+} + Ni(OH)_3 \longrightarrow Ni^{2+} + Co(OH)_3 \quad (28)$$

Precipitation of cobalt as sulfide is done in ammoniacal nickel matte leaching. From sulfate or sulfate/chloride solutions, cobalt is nowadays normally removed by using solvent extraction [46]. Phosphine-based solvent extraction reagents such as di-2-ethylhexylphosphoric acid (D2EHPA) are used to extract cobalt from nickel sulfate solutions. From pure chloride electrolytes, solvent extraction is done with a tertiary amine such as tri-isooctylamine.

Copper may be removed by cementation with activated metallic nickel powder, causing copper ions to precipitate as metallic copper while metallic nickel dissolves (29)

$$Cu^{2+} + Ni(s) \longrightarrow Cu(s) + Ni^{2+} \quad (29)$$

Alternatively, copper may be precipitated as copper sulfide using hydrogen sulfide, sodium sulfide, or ammonium sulfide (30)

$$Cu^{2+} + H_2S(g) \longrightarrow CuS + 2H^+ \quad (30)$$

The MCLE process of Sumitomo uses electrowinning for copper removal. In the copper electrowinning, separation of copper and nickel is achieved by high copper concentration and low pH. High current efficiency of copper removal is achieved by optimizing copper concentration and depositing copper as powder [54]. The Outokumpu process has also used electrowinning to remove copper. Nowadays in the Outokumpu process, copper is leached away before nickel leaching. Solvent extraction is also used to remove copper from the leachate and anolyte of copper-rich anodes.

Iron is normally removed by oxidizing the iron to Fe^{3+} and precipitating it as ferric hydroxide by hydrolysis (31)

$$Fe^{3+} + 3H_2O \longrightarrow Fe(OH)_3 + 3H^+ \quad (31)$$

The pH of the solution is raised by adding, for example, nickel carbonate. Iron hydroxide coprecipitates other impurities from the anolyte, particularly lead and arsenic [43]. Oxidation of iron is accomplished by using air in pure sulfate electrolyte systems and by using chlorine in chloride systems. Iron can be removed from pure chloride electrolytes by solvent extraction with tributylphosphate.

4.1.5 Zinc Production

In 2003, the world zinc production was 9 880 000 tons [55]. The most important zinc production process is the electrolytic or roast-leach-electrowinning (RLE) process. This was first used in 1916 by Anaconda and Cominco. The industrial processes of zinc production use zinc oxides as raw materials. The most important natural raw material is zinc sulfide, and, therefore, it needs to be roasted and converted to oxide. The main problem in leaching and liquor purification is separation of zinc and iron. As dissolution of iron cannot be avoided, it must be precipitated from the zinc sulfate solution. Impurities still present after the iron precipitation stage can lead to lower current efficiency and impurities in the zinc cathode. Therefore, the solution is further purified by cementation with zinc powder.

There are several other process routes to produce zinc-containing liquor for electrowinning. An alternative to the RLE process is the direct leach-electrowinning process. In direct leaching, the zinc concentrate, the regenerated sulfuric acid from subsequent electrowinning, and oxygen are fed into the leaching reactor. In the Sherritt zinc pressure leach process, the zinc concentrate is leached with sulfuric acid and ferric ions at 145–150 °C. During leaching, the sulfide of the mineral is oxidized to elemental sulfur. The sulfuric acid is the spent acid from zinc electrowinning. Ferric sulfate comes from the ferrite-leaching step. The leach liquor pH is raised to 3.5 with zinc hydroxide to remove calcium as gypsum. The solution purification includes copper removal with zinc dust, iron removal by increasing the pH to 5, and removal of the remaining impurities with zinc dust [56]. The Outokumpu direct leach process is done under atmospheric pressure. The temperature is kept at about 100 °C, and sulfuric acid concentration is 10–40 g L^{-1}. Oxygen is fed in the acidic slurry. Sulfur in zinc concentrate is converted to the solid elemental form during leaching, and solution purification and electrowinning can be used as in the RLE process.

Teck Cominco has developed a new integrated process, known as *HydroZinc*, for the recovery of zinc from sulfide ores. The process includes heap leaching, neutralization, solvent extraction, and electrowinning. The zinc sulfide ore is crushed, agglomerated, and stacked on a lined pad for heap leaching. In the heap, bacteria oxidize sulfides to sulfur or sulfate, releasing metals to solution. The heap is aerated from the bottom and irrigated from the top with solvent extraction raffinate that contains approximately 15 g L^{-1} sulfuric acid. Pregnant leach solution that is rich in zinc and low in acid is collected from the bottom of the heap and treated in the neutralization circuit. Zinc is recovered by solvent extraction and electrowinning. The current density ranges from 350 to 640 A m^{-2} [57].

Electrowinning from zinc sulfate solution is the usual final stage in the recovery of zinc. Approximately 80% of total world production is electrowon. Two typical methods for zinc electrolysis were previously used in industry. Either the electrolytes were slightly acidic and current densities were about 325 A m^{-2} or they were strongly acidic and current densities were close to 850 A m^{-2}. The process parameters widely used today are current densities in the range 400–600 A m^{-2} and a process temperature of 30–40 °C [5].

4.1.5.1 Electrodeposition of Zinc

In principle, zinc is more difficult to electrowin than other base metals because the equilibrium potential of zinc is considerably lower than that of hydrogen. The evolution of hydrogen from an acidic zinc sulfate solution is thermodynamically favored over zinc deposition. Zinc has high hydrogen evolution overpotential, and this makes the electrolytic production of zinc possible. The initial deposition of zinc on the fresh cathode surface is influenced by simultaneous zinc deposition and hydrogen evolution. The overpotential must be high enough for zinc deposits to nucleate. Impurities, sometimes even at concentrations as low as the parts per billion level, could decrease hydrogen evolution overpotential and thus partly or entirely stop the electrodeposition of zinc. The electrolyte impurities can be grouped into four groups on the basis of their effect [58]:

1. Ag, Cu, and Cd are capable of codepositing with zinc at their limiting current density. They will induce zinc redissolution by forming a cathodic area.
2. Ni, Co, and Fe codeposit with Zn. They will lower the hydrogen overvoltage and initiate cyclic zinc deposition and redissolution.
3. As, Sb, and Ge act as catalyst for hydrogen evolution. They will increase the critical current density at which cathode formation begins, but once zinc has deposited, they will not cause zinc redissolution.
4. Al and Pb will not cause zinc redissolution.

The hydrogen overvoltage is also influenced by the condition of the surface of the electrodes. The starter sheets must be smooth as smooth surfaces lead to higher hydrogen evolution overvoltages. For this reason, the cathodes must be stripped after one to three days in practical operation.

The electrolyte typically contains 55–70 g L^{-1} Zn, 150–200 g L^{-1} sulfuric acid, 10–20 g L^{-1} magnesium, and 2–5 g L^{-1} manganese. The temperature is 35 to 40 °C. Small, controlled quantities of glue and antimony are added to the electrolyte to help form a smooth deposit. Barium and strontium carbonates are added to minimize Pb contamination of the zinc. The presence of antimony up to 0.08 ppm increases the limiting current density of Zn^{2+}. Antimony enhances hydrogen evolution reaction and decreases the cathode diffusion layer thickness [59]. The electrolyte composition is a compromise of high zinc deposition rate, high hydrogen evolution overpotential, and low resistivity. To achieve a high current efficiency, it is necessary to maintain high electrolyte purity, high $ZnSO_4$ concentration, and low H_2SO_4 concentration, and proper electrolysis conditions including electrode configuration and maintenance, current density, and temperature. In electrowinning, the current efficiency will increase as the Zn/acid ratio increases. This results in a significant increase in power consumption as the zinc concentration falls below 50 g L^{-1} [60].

Zinc deposits grow as hexagonal platelets. Depending on the extent of polarization, electrolyte impurities, and electrolyte additives, the platelets can be basally, randomly, or vertically oriented to the cathode surface. Low overpotential with impurities (Sb, As, Ge, Co, Ni, Cu) produces basal deposits. Intermediate overpotentials with impurities and additives in the electrolyte produce platelets with 30–70° angle to the surface. High overpotentials with organic additives produce platelets that deposit at

70–90° to the cathode surface [61]. Stripping of cathode zinc is more difficult with basal or vertical orientations.

4.1.5.2 Energy Consumption

In electrowinning processes using sulfuric acid solutions, the main anodic reaction is the decomposition of water with evolution of oxygen and formation of hydrogen protons. The largest component of the cell potential is the decomposition voltage of the electrolyte. The polarization of the anode strongly affects the cell potential, because the evolution of oxygen always needs certain overpotential. Table 7 lists factors affecting the cell voltage which is a sum of thermodynamical equilibrium potentials of decomposition of water and zinc deposition, electrode overpotentials, and ohmic potential drops.

The current density is between 400 and 800 A m^{-2} giving the cell voltage 3.3–3.5 V. The energy consumption is about 3.1–3.3 kWh kg^{-1}. The cathodic current efficiency is 88 to 93%, the difference being due to H_2 evolution, short circuits, and current lost to the ground. The anodic current efficiency for oxygen evolution is ~98%, and the difference is due to MnO_2 and PbO_2 formation and Cl_2 evolution. When operated at a current density of 400 to 500 A m^{-2} and a current efficiency of 92–94%, the energy consumption is typically around 3.1 kWh kg^{-1}, which is about 80% of the total electrical power required for zinc production in RLE plants. Over 50% of power costs in the electrowinning are caused by the anode.

Tab. 7 Formation of cell voltage in zinc electrowinning

Potential source	*Overpotential [V]*
Thermodynamical voltage	1.99
Electrolyte resistance	>0.6
Cathode overpotential	~0.1
Anode overpotential	>0.6
Hardware resistance	~0.2
Total	3.5

In practice, only some of the components of cell voltage can be reduced. Thermodynamical potentials will always remain the same. Ohmic potential drops due to the resistivity of the electrolyte and hardware cannot be significantly reduced. The effect of electrolyte resistance on cell voltage is determined by the current density, the conductivity of the electrolyte, and the size of the gap separating the electrodes. High acidity and operating temperature increase the conductivity by improving ion transport across the electrode gap, but high acidity lowers the current efficiency. This leaves only the electrode overpotentials, and only by reducing anode overpotential, significant improvements are possible. The oxygen evolution overpotential on lead anodes is several hundred millivolts, and in zinc electrowinning the amount of anode voltage (the sum of thermodynamical decomposition potential and various overpotentials) is about 50–60% of the cell voltage. Several methods have been attempted to reduce the anode overpotential. The anodic reaction of water decomposition could be replaced by methanol oxidation. The dimensionally stable, coated titanium anodes have lower oxygen evolution overpotential than Pb–Ag anodes. The addition of certain chemicals into the electrolyte can reduce anodic polarization.

4.1.5.3 Zinc Electrowinning Practice

The electrowinning of zinc is done from highly acidic-purified electrolyte. Purification of the solution is done by adding

zinc dust. Zinc will displace more noble elements by cementation reaction. The purified $ZnSO_4$ solution feed to the cellhouse contains 150–175 g L^{-1} Zn. It is diluted by circulating the cell solution. The electrowinning cell concentrations are then 50–60 g L^{-1} Zn and 170–210 g L^{-1} H_2SO_4. The cells are undivided. The anodes are usually lead alloyed with silver and the cathodes are aluminum plates. A typical tankhouse may contain 20 000–35 000 cathodes and anodes with an additional anode in each cell. The area of the electrodes varies from 1.1 to 1.5 m^2, but the trend is to use larger electrodes of area 3.0–3.5 m^2. The electrode spacing is 80–100 mm. The cathodes are grown for 24–48 hours only to prevent dendrite formation. After stripping, the zinc cathodes are melted and cast into ingots.

The quality of zinc deposit depends on the purity of the electrolyte. With a pure electrolyte, it is possible to use higher temperatures, and thereby lower electrolyte resistance and decrease electrode overvoltages. With an impure electrolyte, the temperature must be lowered to 30–35 °C to hinder hydrogen evolution caused by the impurities. Temperatures below 30 °C can cause formation of calcium sulfate; temperatures above 40 °C can increase lead corrosion, and above 45 °C can increase sticking of the deposit.

Demands on the anode material are low corrosion rates for a long lifetime, low oxygen overvoltage to minimize the specific energy consumption, and good mechanical stability. The lead anodes are a good compromise for anodes in zinc electrowinning. Lead is a low-cost material with good corrosion resistance in strong corrosive environments. The anodes used in the zinc electrowinning process are usually conditioned before use. The conditioning process can be passivation in a $KF-H_2SO_4$ electrolysis bath, by chemical oxidation in a $KMnO_4-H_2SO_4$ solution or by mechanical surface roughening. These treatments promote the formation of a PbO_2 layer, which adheres well to the substrate. All three techniques roughen the anode surface to enhance the generation of a hard adherent corrosion layer and shorten its required formation time. Well-conditioned anodes produce Zn cathodes with low Pb contamination within a week of usage in the cellhouse. On the other hand, if the anodes are not conditioned, they corrode excessively and precipitate large quantities of MnO_2 [13]. A very thick MnO_2 layer increases the anode overvoltage and can cause short circuits.

The aluminum cathode sheets are 5–7 mm thick, supported on carrier bars, and provided with copper contacts. The Al_2O_3 layer prevents sticking of zinc. The cathodes are polished at intervals of a few weeks to prevent excessive adhesion of the deposits. Too high a fluoride (over 50 mg L^{-1}) or chloride (300 mg L^{-1}) level pits the aluminum starter sheets, causing zinc to stick very tightly to the starter sheets.

4.1.6 Noble Metals

The refining of noble metals is done mostly using hydrometallurgical processes. The typical raw materials are ores that are treated with complex leaching and base metal refinery by-products. The smelting and converting of copper sulfide concentrates result in blister copper that contains almost all of the noble metals present in the original concentrate. Upon electrolytic refining of copper, insoluble impurities accumulate at the bottom of the refining tank. These slimes contain the noble metals of the concentrate. The slime may

contain, for example, up to 20% silver. The anode slime is collected and smelted in a small furnace to oxidize virtually all metals present except silver, gold, and PGMs. The recovered metal is called *doré*, and it generally contains 0.5–5% gold, 0.1–1% PGMs, and the rest, silver. This metal is cast to anodes and electrorefined in a solution of silver–copper nitrate. Electrolytic refining is used mostly for gold and silver, whereas PGMs are recovered by using precipitation as compounds.

4.1.6.1 Gold

Of the various gold-bearing minerals existing in nature, the most important are native gold, electrum, and tellurides. Gold is often included in sulfide minerals and in selenium and antimony compounds. Native gold is the main form of gold occurring in nature, with a gold content of 90% or more and frequently accompanied by silver. With silver, gold forms an alloy-type compound called *Electrum* which contains 15–35% silver. Copper and iron are the most common impurities in gold ores. Gold-containing raw materials can be leached using cyanide, thiourea, or strong chloride solutions. The same leaching processes that are used for gold ores can be used for recycling of gold-containing electronic scrap. Separation of gold can be done, for example, by using activated carbon followed by chemical or electrochemical precipitation. Electrowinning from leach liquors or eluates can be done using high–surface area cathodes.

The heap leaching of gold with alkaline cyanide solution produces, typically, liquor containing 0.5–5 mg L^{-1} gold. The gold is recovered by using activated carbon, and stripping of activated carbon produces concentrated gold solutions with 50–1000 mg L^{-1} gold from which the gold is electrowon [62]. During the electrowinning of gold in alkaline solutions, the principal anode reaction is oxygen evolution (32)

$$4OH^- \longrightarrow O_2 + 2H_2O + 4e^- \quad (32)$$

The main cathodic reaction is the reduction of oxygen generated in the anodic reaction. In cyanide solutions, the gold deposits by reaction (33)

$$Au(CN)_2^- + e^- \longrightarrow Au + 2CN^- \quad (33)$$

The current efficiency in most gold electrowinning operations is only 5–10%. The majority of current is used in oxygen reduction. As the anode reaction consumes OH^- ions, the local pH lowers. On the cathode, metal is deposited and hydroxide ions are formed. As a net result, the metal concentration and pH of the electrolyte decrease. The decrease of pH is critical especially in cyanide solutions as anode corrosion increases and at pH <9.0, the formation of lethal HCN gas begins [63].

When the dissolution and recovery are done using acidic thiourea solutions, the cathodic reaction is the reduction of gold-thiourea complex (34)

$$Au(CS(NH_2)_2)_2^+ + e^- \longrightarrow Au + 2CS(NH_2)_2 \quad (34)$$

Electrowinning on titanium or stainless steel cathodes yields high current efficiencies for high gold concentration (~1400 ppm). At low concentrations, the current efficiency can be less than 10%, but the recovery of gold is still nearly 100% [64].

Factors that favor gold deposition over hydrogen evolution are high gold concentration in solution, low current density, increase in temperature, and increase in

conductivity. It is possible to operate cells with a low current density while keeping the cell size reasonable by using high–surface area cathodes. A special cell construction for gold recovery is the Zadra cell. The Zadra process is used to recover gold absorbed in active carbon. In the Zadra process, 1–2 wt% sodium hydroxide and 0.5 wt% sodium cyanide aqueous solution at 70–85 °C is circulated between activated carbon and the electrowinning cell. The solution flows up through a stationary bed of loaded carbon. Gold that has been adsorbed on the carbon as a sodium or calcium/gold cyanide ion pair is desorbed from the carbon by a reversal of the adsorption kinetics. The eluate goes to the Zadra cell for recovery. The Zadra cell has a perforated cylinder as cathode compartment, and the compartment is filled with steel wool. The pregnant solution enters the cell through a central feed pipe to the bottom of the cell and flows upward through the cathode compartment distributing radially. The flow rates are low, usually 10 L min^{-1} giving a residence time of 5 minutes. Outside the cathode compartment is a cylindrical stainless steel anode. Gold is recovered from the pregnant strip solution by electrowinning onto steel wool. The deposit morphology depends on electrolysis conditions, and it can be solid, spongeous, or powder. In all cases, the cathode steel mat is sacrificed when the gold is melted [3].

Another cell for gold recovery is the Mintek cell. It has a rectangular configuration with alternating anodes and cathodes in parallel. The anodes are perforated stainless steel plates. The cathodes are removable plastic baskets with perforated sides, packed with steel wool. The electrodes fit the cell sides and bottom, so that the electrolyte has to flow through the electrodes. Packed cathodes of steel wool or stainless steel wire provide a very high cathode surface area while maintaining good cathode porosity. The flow of solution is through the cell, that is, parallel to the flow of current. In Mintek-type cells, the potential drop that occurs across the cathode bed limits cathode thickness. The cathodic overpotential inside the cathode bed decreases as the distance from the anode increases. The overpotential decrease depends on the conductivity of the electrolyte. On the bed surface, the deposition is mass transport controlled, and deeper inside, it is kinetically controlled. If the cathode bed is too thick, gold deposition stops inside the cathode. The thickness of packed bed cathodes is limited to a few centimeters. As the equilibrium potential of gold deposition decreases with decreasing gold concentration, thinner cathodes are needed for more dilute solution to maintain cathode activity.

The first application for gold electrolysis was the refining of impure gold. Emil Wohlwill developed the Wohlwill electrolytic refining process for gold in 1878 at the Norddeutsche Affinerie in Hamburg. It is still in use for the production of fine gold of quality >99.9% on a small scale. The impure gold is in the form of cast anode plates. The electrolyte composition of the Wohlwill process varies, but typical values are 140–200 g L^{-1} gold and 100–150 g L^{-1} hydrochloric acid. The electrolyte must have high acidity to avoid hydrolysis of gold and high gold concentration to get a solid deposit. Electrolysis is carried out with agitation at 65–75 °C and high current density up to 2500 A m^{-2} [65]. The cell voltage depends on current density and it can vary between 0.9 and 2.5 V. The cathodes were originally made of fine gold of thickness 0.25 mm. These have now largely been replaced by titanium cathodes, from which the deposited fine

gold can be stripped. The accumulation of metallic impurities in the electrolyte, and of anode slime in the cell, mean that the electrolyte can only be used for a limited period of time [66].

During electrorefining at the anode, the gold dissolves as Au^{3+} and at the cathode, gold is deposited. Some of the gold dissolves anodically as Au^{+}, but the Au^{+} ions disproportionate immediately and form gold powder and Au^{3+}. Therefore, the anodic current efficiency for Au^{3+} dissolution is over 100%. Platinum and palladium also dissolve and remain in the electrolyte. Silver is dissolved, but it will form insoluble silver chloride and fall to the anode slime. Platinum and palladium can be recovered from the electrolyte, which also contains copper, iron, and nickel. The anode slime is collected. In addition to silver chloride, it contains gold formed in the disproportionation reaction, rhodium, iridium, ruthenium, and osmium, which can be recovered. About 10% of the anode gold, especially parts located above the electrolyte, remains as scrap that is remelted to new anodes. The operation of a Wohlwill cell ties up precious metal, and so the tendency is to use small cells at high current densities.

4.1.6.2 **Silver**

Approximately 85% of world silver production is refined by electrolysis [67]. While the aim in many other metal electrorefining processes is to produce a massive product, the preferred deposit morphology in silver refining is dendrites. This is achieved by depositing the metal onto a permanent stainless steel cathode from an electrolyte without deposit-leveling additives, and operating at high current density. The anodes are cast silver derived, for example, from the slimes formed in electrorefining of copper or lead. The silver dendrites formed on the cathode can be harvested mechanically since they do not attach strongly to the electrode surface. The electrorefined metal contains between 99.9 and 99.99% silver. The anodic slimes from silver refining are further treated to recover gold and PGMs.

The electrolyte typically contains 40–60 g L^{-1} silver and nitric acid to lower the pH to 1–1.5. The cell operates at 35 °C and the current density is 400–650 A m^{-2}. The cell voltage is in the range 2.0–2.5 V. The energy consumption is 0.6 kWh kg^{-1} Ag. To compensate for the anodically dissolved base metals, silver nitrate solution is fed into the electrolyte to prevent depletion of silver. A single electrolysis gives a purity of 99.95–99.99% Ag and a double electrolysis gives a purity of 99.995–99.999%. The slime contains metals more noble than silver (Ag, Pt), detached silver particles, lead dioxide, and copper oxides. Copper accumulates in the electrolyte. If the concentration of copper or selenium becomes too high, they can codeposit with silver [65, 66].

Two different techniques for silver electrorefining are employed, the Möbius and Balbach-Thum cells. The main difference between them is that the electrodes are arranged vertically in the Möbius cell and horizontally in the Balbach-Thum cell (Fig. 22). In the Möbius cell cast, crude silver anodes are suspended in anode bags that collect anode slime. The spent anodes are replaced continuously by new anodes. The method uses stationary or rotating stainless steel cathodes. Deposited silver forms dendritic crystals on the cathode and these are removed manually or continuously by automatic scrapers. The Möbius electrolysis process is most suitable for crude silver containing >90% Ag. With high copper content, the electrolyte must be regenerated very often, and with

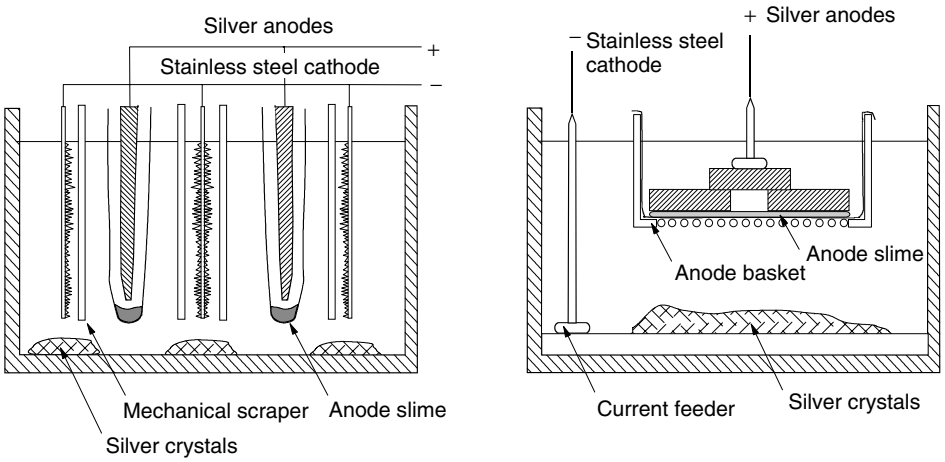


Fig. 22 Möbius and Balbach-Thum cells for silver refining.

high gold content, the anode bags require frequent emptying. In the Balbach-Thum cell, the anodes and cathodes are arranged horizontally. A cell has one anode and one cathode compartment. The cathode is a stainless steel plate covering the entire cell bottom area. Above the cathode is a basket of nonconducting material with fabric bottom containing the anodes. As the silver anodes are completely covered by electrolyte, they dissolve completely. The current is supplied via silver contacts. The anode slime is recovered more easily than in Möbius cell. The disadvantages include the large space requirement and higher energy consumption [65].

The silver used in the photographic industry can be recycled from spent photographic processing solutions and photographic film. The solutions are processed on site electrolytically, while the film is burned and the ashes are leached to extract the silver content [10]. In photographic baths, silver forms various complexes with the anions, such as $S_2O_3^{2-}$, SO_3^{2-}, Br^-, or Cl^-, present in the solution. These complexes are principally $Ag(S_2O_3)_2^{3-}$ and $Ag(S_2O_3)_3^{5-}$. Electrowinning of silver from photographic fixing solutions has been studied from the beginning of the twentieth century and there are many industrial installations. Massive and porous cathodes have been used. Massive cathodes have been principally made of stainless steel or graphite. The main disadvantage of massive electrodes was the final silver concentration that can be larger than the recommended value for discharge, typically 0.1 to 1 mg L^{-1}. Porous, three-dimensional cathodes have been, for example, carbon felt, carbon fibers, grids of expanded stainless steel, and rotating cylinder electrodes. The porous electrodes are at present the most efficient in silver removal from photographic wastes. The major problem encountered is the plugging up of the electrodes [68].

4.1.6.3 Platinum-group Metals

In 2003, world mine production of PGMs was 453 000 kg [69]. The separation of PGMs utilizes their complex chemistry and solubility in *aqua regia*. Platinum, palladium, and gold will dissolve in

aqua regia, whereas ruthenium, rhodium, iridium, and osmium will not [5, p. 832]. The recovery of PGMs from the solution is usually done with chemical precipitation. From chloride solutions, ammonium-chloride complexes can be precipitated. Solvent extraction is also used for PGM recovery [70].

Hydrometallurgy is used in refineries to dissolve the PGMs for final purification. The most desirable process for PGM concentrate processing would dissolve both base metals and PGMs, eliminate the residue, and then separate each with high efficiency and low losses. To achieve this, a suitable chelating agent to hold the metals in solution is needed. The agents might include, for example, chloride and other halides, cyanide, thiocyanate, thiosulfate, thiourea, and ammonia. A significant chemical characteristic of dissolved PGMs is that the ions are strong oxidants, easily reduced to metal, and easily hydrolyzed. Thus, they tend to precipitate and deposit in undesirable locations, and this can make the development of a robust, controllable process difficult [71].

Leaching and electrolysis processes can be used for metal recovery from waste electrical and electronic equipment. Metals such as Ag, Au, Cu, Pb, Pd, Sn, are dissolved from shredded electronic scrap in an acidic aqueous chloride electrolyte by oxidizing them with aqueous dissolved chlorine species. In the electrochemical reactor, chlorine is generated at the anode for use as the oxidant in the leach reactor and the dissolved metals are deposited from the leach solution at the cathode. The very low concentrations of the precious metal ions require the use of porous electrodes with high specific surface areas and high mass transport rates to achieve economically adequate reactor productivities and space–time yields [72].

4.1.7 Other Metals

4.1.7.1 Lead

World production of refined lead was 6.39 million tons in 2002 [73]. Primary lead production has usually two distinct operations: smelting lead concentrate to bullion and refining the bullion. Depending on the nature of the concentrate, the bullion contains impurities, such as iron, copper, nickel, cobalt, zinc, arsenic, antimony, tin, bismuth, selenium, tellurium, silver, and gold, as well as sulfur and oxygen. In conventional lead refining, the bullion from the smelting is drossed by cooling the bullion to within 50 °C of the solidification point of lead. Iron, copper, and many impurities from the bullion are rejected by crystallization partly or fully. For the most part, conventional lead pyrometallurgical refining is conducted in gas-fired steel kettles in batch operations.

The electrolytic refining of lead bullion anodes has been practiced since the beginning of the twentieth century. In practice, the useful solutions have been restricted to the soluble lead salts of fluosilicic acid (H_2SiF_6), fluoboric acid (HBF_4), and sulfamic acid (NH_3SO_3). The most important of these is the fluosilicic acid electrolyte. All these electrolytes require the use of additives for deposit control and the relatively expensive electrolytes cannot be regenerated. The electrolytes and operating parameters are quite similar. The electrolytes contain 80–100 g L^{-1} Pb and 50–70 g L^{-1} free acid at temperature of 30–40 °C. The current densities are between 160 and 200 A m^{-2} and cell voltage 0.3–0.4 V. Current efficiencies are high, 96–98%. The energy consumption is 0.11–0.13 kWh kg^{-1} [41].

The size of the lead anodes is limited owing to the metals' low strength. The

composition of anodes also has some limitations. The copper content must be no more than 0.04–0.05%; or a passivating layer will form. On the other hand, the Bi + Sb content must be over 1% to get a layer that maintains its shape and allows the diffusion of dissolved metals into the electrolyte [5, p. 522]. Metals with a higher electrochemical potential than lead include silver, gold, copper, bismuth, antimony, arsenic, and germanium. They remain essentially undissolved and accumulate in the anode slime. Metals with a lower potential, such as iron, nickel, and zinc dissolve, but as their concentration in the anode is low, their enrichment in the electrolyte is not significant. The major impurity that can codeposit with lead is tin. The equilibrium potential of tin is −0.140 V, which is close to that of lead, −0.126 V. Tin is codeposited with lead from fluosilicate and fluoborate electrolytes, but in the sulfamic acid electrolyte, lead and tin are separated because of the low solubility of tin sulfamate.

The Betts electrorefining process was developed by Cominco at its smelter in Trail, British Columbia, Canada, in 1902. Lead from the smelter is melted and cast into anode slabs. Cathode starter sheets are made from refined lead. The electrolyte composition varies from 60 to 100 g L^{-1} lead as lead hydrofluorosilicate, 70–140 g L^{-1} free hydrofluorosilicic acid, and 1–1.25 g bone glue for ton of refined lead to help form a smooth deposit. The consumption of hydrofluorosilicic acid is around 2 kg t^{-1} of electrolytic lead. The process operates at 30 to 40 °C; higher temperatures decrease specific energy consumption but cause dendritic growth and higher evaporation of HF from the electrolyte. A current density between 160 and 200 A m^{-2} is used and the current efficiency is 85–95%. The cell voltage is 0.35 to 0.65 V and the energy consumption is 0.15–0.19 kWh kg^{-1}. The cell voltage rises during operation with increasing thickness of the anode slime layer and may reach up to twice the initial voltage. The energy consumption imposes an economic limitation for increasing current density. In addition, the increasing voltage can lead to the dissolution of antimony, arsenic, and bismuth from the slime. The residence time of the anodes in the cell is governed by this rise in cell voltage and thus by the concentration of impurities in the anode lead [74].

In the Betts process, the lead content of the electrolyte rises slowly because anodic current efficiency is higher than the cathodic one. The lead content is usually maintained at the desired level by electrolysis in separate cells with insoluble graphite anodes. The cathodic reaction is (35)

$$PbSiF_6 + 2H^+ + 2e^- \longrightarrow H_2SiF_6 + Pb \quad (35)$$

The removal of lead from the electrolyte is an identical process to decopperizing copper electrorefining electrolyte bleed in ''liberator'' cells.

Fluoborate electrolytes have been tested for the production of lead using divided cells. HBF_4 has a number of technical advantages over H_2SiF_6, such as lower HF vapor pressure, higher solubility for Pb, and higher conductivity. The applications have been the electrochemical recovery of Pb from lead concentrates, refining of lead bullion, and recovery from recycled lead-acid batteries [75]. The anodic reaction is the oxidation of ferrous ion to ferric ion on graphite electrodes (36)

$$2Fe(BF_4)_2 + 2BF_4^- = 2Fe(BF_4)_3 + 2e^- \quad (36)$$

The cathodic reaction is the deposition of lead on stainless steel electrodes (37)

$$Pb(BF_4)_2 + 2e^- = Pb + 2BF_4^- \quad (37)$$

The process has been run on industrial scale for lead concentrates [75].

The PLACID electrowinning cell has been developed to treat lead-bearing dusts, slags, and spent lead-acid batteries. The raw materials are leached in warm, slightly acidic chloride solution to form soluble lead chloride. The impurities are removed by cementation with lead powder. The purified solution is electrolyzed, lead is deposited as sponge, and hydrochloric acid generated on the anode is collected and returned to leaching. The cell uses DSA-type anodes, titanium bar or mesh cathodes, and the cell is divided using cation-selective membrane. The lead sponge is removed mechanically from the cathodes and collected from the cell bottom using a belt conveyor. The electrolyte contains 10–12 g L^{-1} $PbCl_2$, 4–5 g L^{-1} HCl and 260 g L^{-1} NaCl. The temperature is 60–70 °C. The operating current density is 1000–1400 A m^{-2}. The anodic current efficiency is >95% and the cathodic one is >99%. In industrial scale tests, the cell voltage was 3.5 V and energy consumption was 0.9 kWh kg^{-1} lead [76].

4.1.7.2 Cobalt

Cobalt is produced as a coproduct of nickel or copper refining. Copper–cobalt sulfide concentrates can be processed by the RLE process. Mixed cobalt–nickel sulfides can be precipitated from ammoniacal leach solutions and as mixed nickel–cobalt hydroxide or carbonate from acid sulfate leach processes. From chloride leach solutions, cobalt can be separated by solvent extraction. Most cobalt production is associated with nickel production from sulfide and laterite ores. Pressure leaching, solvent extraction followed by the electrowinning of nickel and cobalt, is the preferred process, replacing the reduction roasting and electrorefining process route. Today, most of the world's cobalt is produced by electrowinning [46, 77]. In 2003, the world production was 40 300 tons of cobalt metal, metal powder, and compounds [78].

Many processes utilize electrowinning for cobalt recovery. Cobalt can be deposited as high-purity metal from sulfate or chloride solutions. The electrowinning solution must be purified from metallic impurities. The following methods have been used [5, p. 820]:

- Removal of copper by cementation with cobalt, deposition as sulfide using H_2S, solvent extraction, and electrowinning.
- Removal of iron by oxidation with air to Fe^{3+} followed by hydroxide precipitation.
- Removal of arsenic by crystallization as As_2O_3 or deposition with iron as arsenate.
- Removal of nickel by deposition as amine-sulfate double salt or by solvent extraction.

After solution purification, cobalt can be deposited as Co(II) or Co(III) hydroxide. The cobalt solutions are often too low in cobalt for direct electrowinning. In this case, cobalt hydroxide is precipitated, thickened, and redissolved to cobalt tankhouse electrolyte and then electrowon [77]. The cobalt hydroxide has been reduced in electric furnaces to an alloy with 80% Co, 18% Fe, and 0.8% Ni, cast to anodes, and electrorefined. The process was done in Hybinette cells, and the solution contained 60–120 g L^{-1} Co, pH was 3.3–48, and the temperature was 60 °C. The current density was 160–240 A m^{-2}, cell voltage was 1.2–1.8 V, and the current efficiency was 95%. This process is not used anymore because of the lack of

suitable raw materials. Instead, cobalt is recovered by electrowinning and hydrogen reduction [5, p. 820].

The standard electrochemical potential of cobalt is slightly lower than that of nickel, so nickel should be preferentially reduced during electrowinning. The polarization of cobalt is smaller than that of nickel so kinetically cobalt deposition is favored over nickel. A good current efficiency is obtained at pH of 5.5–6.5 and so the catholyte is separated by a diaphragm from the acid developed in sulfate electrolysis anolyte. Hydrogen evolution accompanied by cobalt deposition results in reduced current efficiency and higher power costs, an increase in pH at the cathode surface, which in turn can cause cobalt hydroxide precipitation, hydrogen incorporation in the deposit that may cause hydrogen embrittlement, and the formation of streak marks along the cathode surface [79]. The electrowinning can be done in cobalt concentrations from 20 to 80 g L^{-1} Co. The temperature is 50–60 °C. The current density is 250–500 A m^{-2}, cell voltage 4–5 V and current efficiency 85–95% [5, p. 820].

The most important method in cobalt production is RLE that is used for copper–cobalt concentrates in the Democratic Republic of Congo and Zambia. These plants use sulfuric acid leaching followed by copper electrowinning. A portion of the copper electrowinning bleed is used for cobalt recovery. Cobalt is precipitated as hydroxide and redissolved for electrowinning. The hydroxide cake is repulped to a solution of pH 7.0–7.2. The cobalt concentration in electrowinning is 25–45 g L^{-1} Co. The pulp is electrolyzed in air-agitated cells. The pulp reacts with acid formed on the anodes, thereby maintaining the pH sufficiently high. The cobalt is deposited onto stainless steel cathodes. Some plants use anodes made from cobalt alloy containing 82% Co, 14% Si, and 4% Mn [46, 77].

Large amounts of cobalt are produced by chloride hydrometallurgy. The large nickel-producing companies, Falconbridge in Norway and Sumitomo in Japan, have similar processes for cobalt production [77].

4.1.7.3 Tin

Tin production in 2003 was 279 000 tons, of which 11 900 tons was from secondary sources [80]. Today, most tin is used as a protective coating or as an alloy with other metals. Tin is used as a coating for steel cans, in solders for joining pipes or electrical conductors, and in bearing alloys. The common method for tin production is reduction of cassiterite (SnO_2) with carbon or carbon dioxide. The reduction smelting is followed by fire refining of the molten tin to remove iron, arsenic, antimony, and copper. Even though the pyrometallurgical treatment is fairly complicated, it is used more often, as electrorefining is a slow process. Very impure tin concentrates are refined by electrolysis after smelting because fire refining is not efficient for producing high-purity metal from low-grade raw material.

The high electrochemical equivalent of Sn^{2+} favors the use of an electrometallurgical refining process. The main impurities Au, Ag, Cu, Bi, As, and Sb do not dissolve under electrorefining conditions, but will remain in the anode slime. The other impurities Ni, Fe, Zn, and Al can be removed by a pyrometallurgical refining operation. Only lead lies close to tin in the electrochemical series. There are practical problems as simple and inexpensive sulfate or chloride electrolytes lead to spongy or dendritic deposits, and the deposit morphology can be changed only by large additions of additives. Therefore,

the process can only be operated at low current densities, leading to low production rates and low current efficiency. The large amount of expensive metal tied up in the process is economically undesirable [81, 82].

Electrorefining can be carried out in acidic or alkaline medium. The acid electrolytes consist of sulfuric acid and stannous sulfate, with additives such as creosulfonic or phenolsulfonic acids and glue to modify deposit formation on the cathodes. The alkaline electrolytes consist of potassium or sodium stannite and free alkali. When compact cathode deposits are required, alkaline electrolytes are inferior to acid sulfate or halogen solutions in terms of electric energy consumption, productivity, cathode morphology, and operating temperature [82, 83].

The anodes for acid electrorefining contain 94–96% tin, and the cathodes are tin starting sheets, but permanent stainless steel cathodes are also being studied. Current densities in acid electrorefining are about 100 A m^{-2}. The main anode impurity is usually lead. Lead is precipitated as lead sulfate in slimes that form on the anode surfaces. The anode slimes adhere to the anode, and the growth of the slime layer increases the cell voltage and slows the tin dissolution. The anode surfaces must be mechanically cleaned to maintain the anode dissolution rate. The cleaning frequency depends on the anode impurity level [82]. The current efficiency is largely determined by the rate of removal of anode slimes. The energy consumption in acid electrorefining is 0.15–0.20 kWh kg^{-1} tin. Because iron accumulates in the electrolyte, regeneration of the electrolyte is necessary [81]. In alkaline NaOH or Na_2S electrolytes, more impure anodes, with 75% Sn, can be used than those used in an acid medium. A smooth deposit can be obtained without additives. However, current efficiencies are very low, and the process often must be carried out at 90 °C [81].

A special application for alkaline electrolysis is detinning of tinplate. Tinplate, a coating of tin over a steel or iron sheet, is a widely used packaging material. The plating can be done either by an electrolytic or hot dipping process. The thickness of the steel substrate is usually in the range of 0.15–0.55 mm. Because of improved coating technology, the tin content has gradually decreased and the average tin layer has a thickness of 0.2–0.3 μm. Extraction of tin from tinned plates serves the double purpose of preparing steel scrap for remelting and producing secondary tin. In sodium hydroxide solutions, tin dissolves selectively from the steel substrate, mainly as stannate, and it is further oxidized to stannite. Compact tin deposits can be produced from alkaline electrolytes if the Sn^{2+} concentration is kept below 0.5–0.75 g L^{-1}, but tin sponge is quite suitable for tin extraction. The current density must be carefully controlled to ensure primary Sn^{2+} formation at the anode, to obtain satisfactory deposits, and to achieve 75–85% current efficiency at the cathode [83].

4.1.7.4 Cadmium

Cadmium production is often connected with zinc production. The impure cadmium metal produced in the purification of zinc electrowinning electrolyte is dissolved in 150 g L^{-1} sulfuric acid solution and electrowon in a process similar to that for zinc. The impure metals that would disturb electrowinning, for example, Cu, Ni, As, are removed by cementation with impure cadmium. The solution is neutralized with lime. The electrolyte contains 140–180 g L^{-1} Cd, 30–40 g L^{-1} Zn, some chlorides, and animal glue as additives. The electrolysis is done using

lead anodes and aluminum cathodes. The cell voltage in cadmium electrowinning is 2.6–2.9 V. The energy consumption is about 1.5 kWh kg^{-1} cadmium. After separation of thallium with ammonium chloride, the purity of electrolytic cadmium is 99.99% [41].

The cathodes are either stationary plates or rotating drums with aluminum strips. When using rotating cathodes the anodes are positioned around the cathode. The use of the rotating cathode allows four to eight times higher current densities (250–330 A m^{-2}) and four times longer cathode periods. With rotating cathodes, the deposit can be grown for four days instead of one day with stationary cathodes [41]. Too long deposition periods will cause strong dendrite formation and short circuits [84].

References

1. M. Pourbaix, *Atlas of Electrochemical Equilibria*, Pergamon Press, Oxford, 1966, p. 644.
2. T. J. O'Keefe, *J. Electroanal. Chem.* **1984**, *168*(1–2), 131–146.
3. D. Pletcher, F. C. Walsh, *Industrial Electrochemistry*, Chapman & Hall, London, 1990, p. 653.
4. O. Kardos, *Plating* **1974**, *61*(2), 129–137.
5. F. Pawlek, *Metallhüttenkunde*, Walter de Gruyter, Berlin, 1983, p. 865.
6. H. Fischer, *Electrochim. Acta* **1960**, *2*, 50–96.
7. H. Fischer, *Elektrolytische Abscheidung und Elektrokristallisation von Metallen*, Springer, Berlin, 1954, p. 729.
8. R. Winand, *Hydrometallurgy* **1992**, *29*(1–3), 567–598.
9. M. Dubrovsky, T. Huh, J. W. Evans, Fluidized bed electrowinning of metals – a review in *Hydrometallurgy, Research, Development and Plant Practice* (Eds.: K. Osseo-Asare, J. D. Miller), The Metallurgical Society of AIME, Atlanta, 1983, pp. 759–772.
10. F. C. Walsh, *Pure Appl. Chem.* **2001**, *73*(12), 1819–1837.
11. D. Landolt, *J. Electrochem. Soc.* **2002**, *149*(3), S9–S20.
12. R. R. Moskalyk, A. Alfantazi, A. S. Tombalakian et al., *Miner. Eng.* **1999**, *12*(1), 65–73.
13. A. Siegmund, D. Prengaman, Zinc electrowinning using novel rolled Pb-Ag-Ca anodes, *Hydrometallurgy 2003: 5th International Symposium*, Honoring Professor Ian M. Ritchie, Minerals, Metals and Materials Society (TMS), Vancouver, Canada, 2003, pp. 1279–1288.
14. H. B. Beer, Improvements in or relating to electrodes for electrolysis. Br. Pat. 1,147,442. 1965, p. 10.
15. H. B. Beer, Electrode having platinum metal oxide coating thereon, and method of use thereof. U.S. Pat. 3,711,385. 1970, p. 16.
16. M. Moats, K. Hardee, C. Brown, *JOM* **2003**, *55*(7), 46–48.
17. D. L. Edelstein, Copper. U.S. Geological Survey, *Minerals Yearbook 2002*, 2004. On-line, http://minerals.usgs.gov/minerals/pubs/commodity/copper/.
18. P. J. Mackey, A. E. Wraith, *Miner. Process. Extractive Metall.* **2004**, *113*, C25–C37.
19. P. J. Bartos, *Resour. Policy* **2002**, *28*, 85–94.
20. T. Robinson, J. Quinn, W. Davenport, Electrolytic copper refining – 2003 world tankhouse operating data in *Copper 2003 – Cobre 2003*, (Eds.: J. E. Dutrizac, C. G. Clement), Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 3–66.
21. M. Stelter, H. Bombach, *Erzmetall* **2001**, *54*(9), 432–438.
22. T. Marttila, L. Palmu, H. Virtanen, Advanced tankhouse systems by Outokumpu in *Copper 2003*, (Eds.: J. E. Dutrizac, C. G. Clement), Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 121–135.
23. M. P. Amsden, R. M. Sweetin, D. G. Treilhard, *J. Met.* **1978**, *30*(7), 16–26.
24. K., Rinne, Moderne Entwicklungen auf dem Gebiet der Kupferraffinationselektrolyse, *Elektrolyse der Nichteisenmetalle*, Ed. Verlag Chemie, Weinheim, 1982, pp. 67–92.
25. J. E. Hoffmann, *JOM* **2004**, *56*(7), 30–33.
26. P. Stantke, *JOM* **2002**, *54*(4), 19–22.
27. O. Forsén, The role of organic additives on the quality of copper in electrorefining process, *Copper '90. Refining, Fabrication, Markets*, The Institute of Metals, Västerås, Sweden, 1990, pp. 189–197.
28. N. L. Piret, *JOM* **2002**, *54*(4), 17–18.

29. M. Stelter, H. Bombach, Investigations on copper electrorefining at high current densities, *Copper 2003 – Cobre 2003*, Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 555–567.
30. T. Robinson, J. Jenkins, S. Rasmussen et al., Copper electrowinning – 2003 world tankhouse operating data, *Copper 2003 – Cobre 2003*, Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 421–472.
31. J. Peacey, X. J. Guo, E. Robles, Copper hydrometallurgy. Current status, preliminary economics, future direction and positioning versus smelting, *Copper 2003 – Cobre 2003*, Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 205–222.
32. D. E. G. Maschmeyer, P. Kawulka, E. F. G. Milner et al., *J. Met.* **1978**, *30*(7), 27–31.
33. D. S. Flett, *CIM Bull.* **2002**, *95*(1065), 95–103.
34. H. K. Lin, X. J. Wu, P. D. Rao, *JOM* **1991**, *43*(8), 60–65.
35. J. Hoffmann, *JOM* **1991**, *43*(8), 48–49.
36. G. E. Atwood, R. W. Livingston, *Erzmetall* **1980**, *33*, 251–255.
37. R. F. Dalton, G. Diaz, R. Price et al., *JOM* **1991**, *43*(8), 51–56.
38. J. Moyes, J. Lam, V. Santos, The intec copper process: development of a new electrowinning cell in *Copper 2003 – Cobre 2003*, (Eds.: J. E. Dutrizac, C. G. Clement), Canadian Institute of Mining, Metallurgy and Petroleum, Santiago, Chile, 2003, pp. 487–508.
39. P. H. Kuck, Nickel. U.S. Geological Survey, *Minerals Yearbook 2002*, 2004. on-line, http://minerals.usgs.gov/minerals/pubs/commodity/nickel/.
40. D. J. MacKinnon, *CIM Bull.* **1989**, *82*(927), 65–82.
41. R. Kammel, Nichteisen-Schwermetalle in *Chemische Technologie, Band 6, Metallurgie* (Ed.: K. Winnacker, L. Küchler), Carl Hansen Verlag, München, 1973, pp. 215–367.
42. A. I. Levin, V. M. Gabov, *Sov. J. Non-ferr. Met.* **1968**, *9*(1), 40–44.
43. J. R. Boldt Jr., *The Winning of Nickel*, Methuen & Company, London, 1967, p. 487.
44. D. K. E. Kerfoot, Nickel, *Ullmann's Encyclopedia of Industrial Chemistry*, Ed. John Wiley & Sons, Weinheim, 2000.
45. L. S. Renzoni, R. C. McQuire, W. V. Barker, *J. Met.* **1958**, *10*(6), 414–418.
46. D. K. E. Kerfoot, D. R. Weir, The hydro and electrometallurgy of nickel and cobalt in *Extractive Metallurgy of Nickel and Cobalt* (Eds.: G. P. Tyroler, C. A. Landolt), TMS, Phoenix, 1988, pp. 241–267.
47. V. A. Ettel, *CIM Bull.* **1977**, *70*(783), 179–187.
48. R. Wu, M. Oliazadeh, A. M. Alfantazi, *J. Appl. Electrochem.* **2003**, *33*, 1043–1047.
49. Saarinen, H., Seilo, M., Die Nickelelektrolyse nach dem Outokumpu-Verfahren, *Elektrolyse der Nichteisenmetalle*, Verlag Chemie, Weinheim, 1982, pp. 143–162.
50. P. Koskinen, M. Virtanen, H. Eerola, Integrated nickel production in Outokumpu Oy in *Extractive Metallurgy of Nickel and Cobalt* (Eds.: G. P. Tyroler, C. A. Landolt), TMS, Phoenix, 1988, pp. 355–371.
51. K. Knuutila, S.-E. Hultholm, B. Saxén et al., New nickel process increasing production at Outokumpu Harjavalta Metals oy, Finland. *ALTA 1997 Nickel/Cobalt Pressure Leaching & Hydrometallurgy*, 19–20 May 1997, Perth, 1997, p. 20.
52. E. O. Stensholt, H. Zachariasen, J. H. Lund, *Trans. Inst. Min. Metall.* **1986**, *5*, C10–C16.
53. E. O. Stensholt, O. M. Dotterud, E. E. Henriksen et al., *CIM Bull.* **2001**, *94*(1051), 101–104.
54. S. Makino, N. Kemori, N. Matsumoto, Removal of copper and other minor elements in the MCLE process in *Hydrometallurgy of Nickel and Cobalt*, 17–20 August 1997, (Eds.: W. C. Cooper, I. Mihaylov), CIM, Sudbury, 1997, pp. 123–135.
55. J. Plachy, Zinc. U.S. Geological Survey, *Minerals Yearbook 2003*, 2004. on-line, http://minerals.usgs.gov/minerals/pubs/commodity/zinc/.
56. M. J. Collins, E. J. McConaghy, R. F. Stauffer et al., *JOM* **1994**, *46*(4), 51–58.
57. H. L. Lizama, J. R. Harlamovs, S. Bélanger, The Teck Cominco Hydrozinc™ process in *Hydrometallurgy 2003 – Fifth International Conference in Honor of Professor Ian Ritchie*, Vancouver, (Eds.: C. A. Young, A. M. Alfantazi, C. G. Anderson et al.), TMS (The Minerals, Metals & Materials Society), Vancouver, Canada, 2003, pp. 1503–1516.
58. T. Ohgai, H. Fukushima, T. Akiyama et al., *Metall. Rev.* **1996**, *14*(2), 24–39.
59. A. Y. Hosny, T. J. O'Keefe, J. W. Johnson et al., *J. Appl. Electrochem.* **1991**, *21*(9), 785–792.

60. O. M. G. Newman, P. A. Adcock, M. J. Meere et al., Investigations of spouted bed electrowinning for the zinc industry, *Hydrometallurgy 2003: 5th International Symposium*, Honoring Professor Ian M. Ritchie, Minerals, Metals and Materials Society (TMS), Vancouver, Canada, 2003, pp. 1313–1330.
61. R. C. Kerby, H. E. Jackson, T. J. O'Keefe et al., *Metall. Trans. B* **1977**, *8*(4), 661–668.
62. N. P. Brandon, M. N. Mahmood, P. W. Page et al., *Hydrometallurgy* **1987**, *18*, 305–319.
63. G. Deschênes, R. M. Morrison, J. Dionne et al., *Hydrometallurgy* **1992**, *30*, 307–325.
64. C. M. Juarez, A. J. B. Dutra, An investigation on gold electrowinniig from acidic thiourea solutions in *Aqueous Electrotechnologies: Progress in Theory and Practice*, (Ed.: D. B. Dreisinger), TMS, Orlando, 1997, pp. 231–245.
65. G. Cornelius, Die Raffination von Gold und Silber durch Elektrolyse, *Elektrolyse der Nichteisenmetalle*, Verlag Chemie, Weinheim, 1982, pp. 215–226.
66. J. Schatz, Die Edelmetalle in *Chemische Technologie, Band 6, Metallurgie* (Ed.: K. Winnacker, L. Küchler), Carl Hansen Verlag, München, 1973, pp. 519–548.
67. A. Prior, High speed silver electrolysis – a bulk material system with hitherto unexplored possibilities, *EMC 2001: European Metallurgical Conference*, Friedrichshafen, Germany. 2001, pp. 347–354.
68. M. Chatelut, E. Gobert, O. Vittori, *Hydrometallurgy* **2000**, *54*, 79–90.
69. H. E. Hilliard, Platinum-group metals. U.S. Geological Survey, *Minerals Yearbook 2003*, 2004. on-line, http://minerals.usgs.gov/minerals/pubs/commodity/platinum/.
70. G. Beck, H.-H. Beyer, W. Gerhartz et al., (Eds.) *Edelmetall-Taschenbuch*, Hüthig, Heidelberg, 1995, p. 676.
71. J. Milbourne, M. Tomlinson, L. Gormely, Use of hydrometallurgy in direct processing of base metal/PGM concentrates, *Hydrometallurgy 2003 – Fifth International Conference in Honor of Professor Ian Ritchie*, Minerals, Metals and Materials Society (TMS), Vancouver, Canada, 2003, pp. 617–630.
72. D. Pilone, G. H. Kelsall, Metal recovery from electronic scrap by leaching and electrowinning IV, *Hydrometallurgy 2003: 5th International Symposium*, Honoring Professor Ian M. Ritchie, Minerals, Metals and Materials Society (TMS), Vancouver, Canada, 2003, pp. 1565–1575.
73. G. R. Smith, Lead. U.S. Geological Survey, *Minerals Yearbook 2003*, 2004. on-line, http://minerals.usgs.gov/minerals/pubs/commodity/lead/.
74. G. Brecka, K. Hein, H. J. Lange et al., *JOM* **1997**, *49*(4), 62–64.
75. M. Olper, M. Maccagni, Electrowinning and electrorefining of base metals using unconventional electrolytes. *Electrometallurgy*, Toronto, Canada, 2001, pp. 17–32.
76. C. Frias, M. A. Garcia, G. Diaz, Industrial size ''PLACID'' electrowinning cell in *Aqueous Electrotechnologies: Progress in Theory and Practice* 9–13 February 1997, (Ed.: D. B. Dreisinger), TMS, Orlando, 1997, pp. 101–113.
77. L. R. Verney, Overview of cobalt supply and production in *Extractive Metallurgy of Nickel and Cobalt*, 25–28 January 1988, (Eds.: G. P. Tyroler, C. A. Landolt), TMS, Phoenix, 1988, pp. 135–186.
78. K. B. Shedd, Cobalt. U.S. Geological Survey, *Minerals Yearbook 2002*, 2004. on-line, http://minerals.usgs.gov/minerals/pubs/commodity/cobalt/.
79. M. I. Jeffrey, W. L. Choo, P. L. Breuer, *Miner. Eng.* **2000**, *13*(12), 1231–1241.
80. T. Carlin, Tin. U.S. Geological survey, *Minerals Yearbook 2003*, 2004. http://minerals.usgs.gov/minerals/pubs/commodity/tin/.
81. G. G. Graf, Tin, tin alloys, and tin compounds, *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag, 2000.
82. P. von Krüger, E. L. de Souza, Fundamental and applied research on tin electrorefining employing stainless steel cathodes in *Yazawa International Symposium on Metallurgical and Materials Processing: Principles and Technologies*, 2–6 March 2003, (Eds.: F. Kongoli, K. Itagaki, C. Yamauchi et al.), TMS, San Diego, 2003, pp. 397–404.
83. T. Kekesi, T. I. Török, G. Kabelik, *Hydrometallurgy* **2000**, *55*(2), 213–222.
84. W. Schmidt, Die Kadmiumgewinnung bei der Ruhr-Zink GmbH in Datteln, *Elektrolyse der Nichteisenmetalle*, Verlag Chemie, Weinheim, 1982, pp. 163–168.

4.2
The Electrolytic Production of Aluminum

James William Evans
University of California, Berkeley, CA, USA

Donald Paul Ziegler
Alcoa Technical Center, Alcoa Center, PA, USA

4.2.1
The Economic Significance of Aluminum

The aluminum industry is one of the major industries of the United States and several other countries. Figure 1, prepared from the statistics of the International Aluminum Institute [1], shows that North American primary aluminum production grew through the 1990s but has shown a decline in recent years, mostly due to smelter closures in the United States.

This primary production is supplemented by recycling of consumer scrap that is roughly one fifth of consumption. Much of the primary production of aluminum in the United States (about 40% of the total) occurs in the west (Oregon, Washington, and Montana). The US Geological Survey estimates that the US industry was operating at 87% of nominal capacity in 2000. Figures for other countries are also high with Canada at 103%, Russia at 100%, Australia 98%, and China at 98%. Electrical energy is a major cost in the production of aluminum and variations in electrical energy cost have a significant impact on world trade. Table 1, taken from a paper by one of the authors (Evans [2]) shows electricity costs and consumption for primary production in several nations in 1992 (and until very recently there will have been relatively small changes in these numbers). Although energy consumption per tonne of aluminum produced shows little variation, the cost of that energy differs significantly and was least in Canada. Canada is therefore the major exporter of aluminum to the United States (61% of aluminum imports for 1996–1999) with Russia second (18%).

Major uses of aluminum are in transportation (37% of US consumption in 2000), packaging, including beverage cans (22%), and building (15%).

4.2.2
The Production of Aluminum

There have been sporadic attempts to produce aluminum by carbothermic reduction [3, 4]. In this approach, akin to the way iron oxides are reduced to iron in the iron blast furnace, the consumption of electrical energy is avoided or at least reduced. There have also been investigations of the production of aluminum by electrolysis of aluminum compounds other than the oxide (e.g. [5]). Some of these alternative electrolytic technologies have even reached a commercial scale [6] but the only method for aluminum production in industrial use today appears to be electrolysis in Hall–Héroult cells. Consequently, the present paper is confined to these cells. The literature on these cells is large. A recent search of the web of science with the subject ''Hall cell'' and similar subjects revealed 79 titles; ''aluminum electrolysis'' yielded 109 publications. This number excludes papers published in the annual ''Light Metals'' volume of the Minerals Metals and Materials Society (TMS). Light Metals contains approximately forty papers each year on Hall cells. Consequently, the authors have made no attempt at a comprehensive examination of the literature on these topics. Rather we have included

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

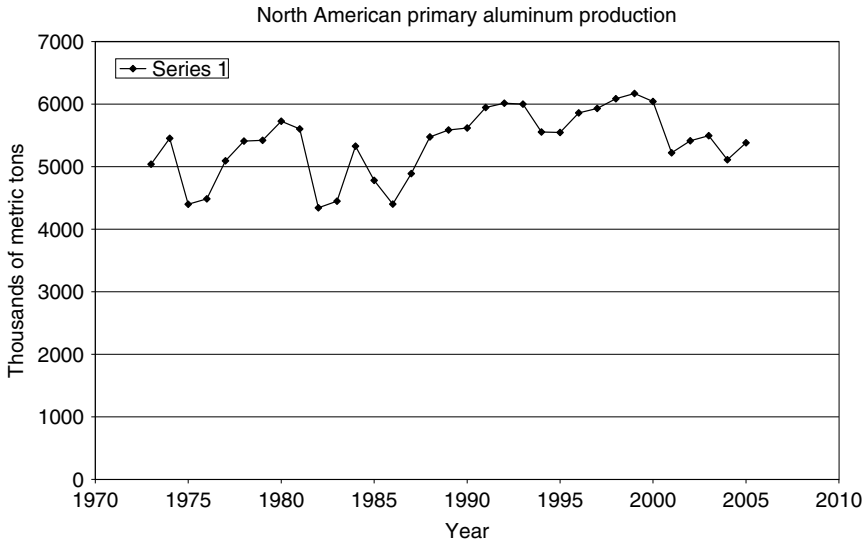


Fig. 1 Primary aluminum production in North America from statistics of the International Aluminum Association [1].

Tab. 1 Average power costs and energy consumption per tonne for primary aluminum smelters in 1992

	Cents [kWh[a]]	*MWh/long[b] [ton]*
Brazil	1.89	16.1
Venezuela	1.15	15.7
United States	2.10	15.8
Canada	1.05	15.6
France	1.96	14.7
Germany	3.03	16.0
Norway	1.70	16.6
Spain	3.40	16.0
Oceania	1.29	15.7

[a]Cents/kWh spread = 2.35cents = 124% of median.
[b]MWh/long ton
spread = 1.9MWh/long ton = 12% of median.
Source: Resources Strategies Inc., Exton, PA 19341.

papers that are known to us and are readily accessible. We will have failed to mention many fine works and for this we apologize.

Good descriptions of the production of aluminum can be found in the literature (Grjotheim et al. [7], Grjotheim and Welch [8], Grjotheim and Kvande [9], Burkin [10], and Peterson and Miller [11]). Referring to Fig. 2 [12], the first step in the production of aluminum from its ore (''bauxite'') is the selective leaching of the aluminum content (present as oxides/hydroxides of aluminum) into hot concentrated NaOH solution to form sodium aluminate in solution. After solution purification, very pure aluminum hydroxide is precipitated from the cooled, diluted solution by addition of seed particles to nucleate the precipitation. After solid–liquid separation the alumina is dried and calcined. These operations are the heart of the Bayer process and the alumina produced is shipped to a smelter where the alumina, dissolved in a molten salt electrolyte, is electrolytically reduced to liquid aluminum in Hall–Héroult cells. This liquid aluminum,

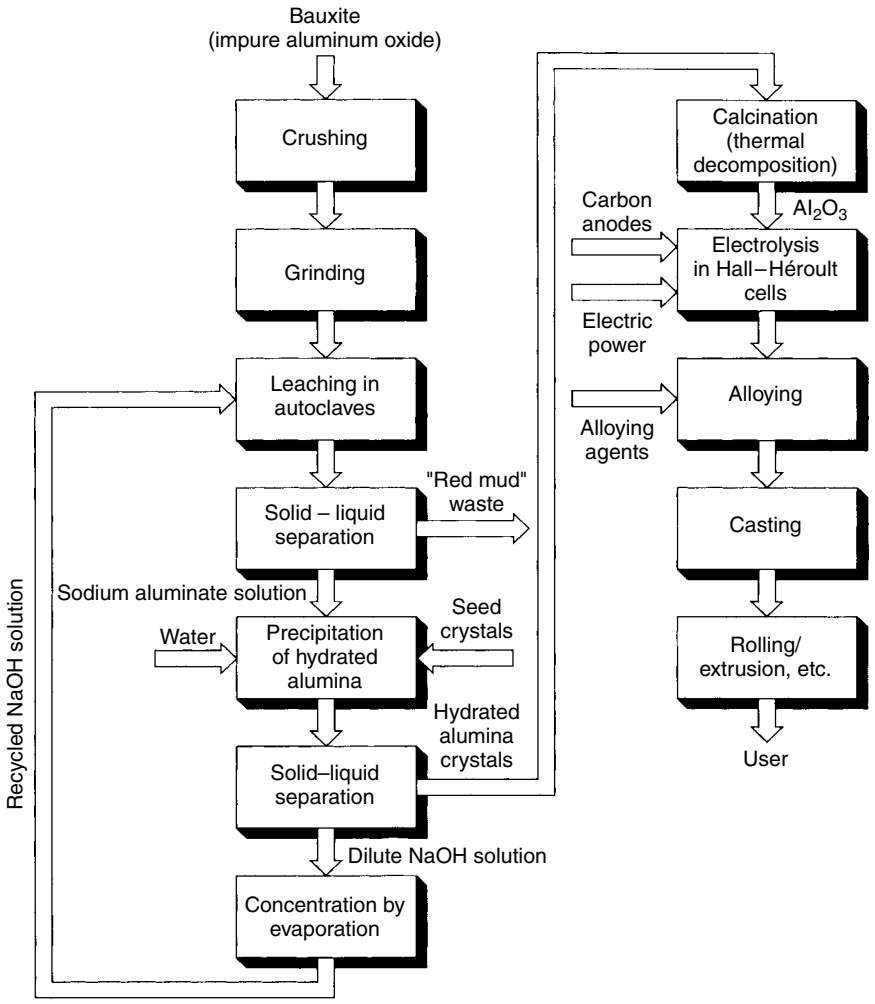


Fig. 2 Schematic diagram of the steps entailed in producing aluminum from bauxite. From Evans and DeJonghe [12].

sometimes after an intermediate step to eliminate dissolved trace impurities (e.g. alkali metals) is alloyed and cast, usually in semicontinuous casters producing ingot (for rolling) or billet (for extrusion), although continuous casters producing strip are gaining ground. A significant amount of aluminum production is from recycled metal and this recycled metal, after melting, is treated by sparging reactive gasses, to remove minor constituents (e.g. magnesium) in the recycled alloy, and by filtration, to remove particles.

4.2.3 The Hall–Héroult Cell

Figure 3, from McGravie et al. [13] shows, in cross section, the three types of Hall–Héroult cell. The cross section of the

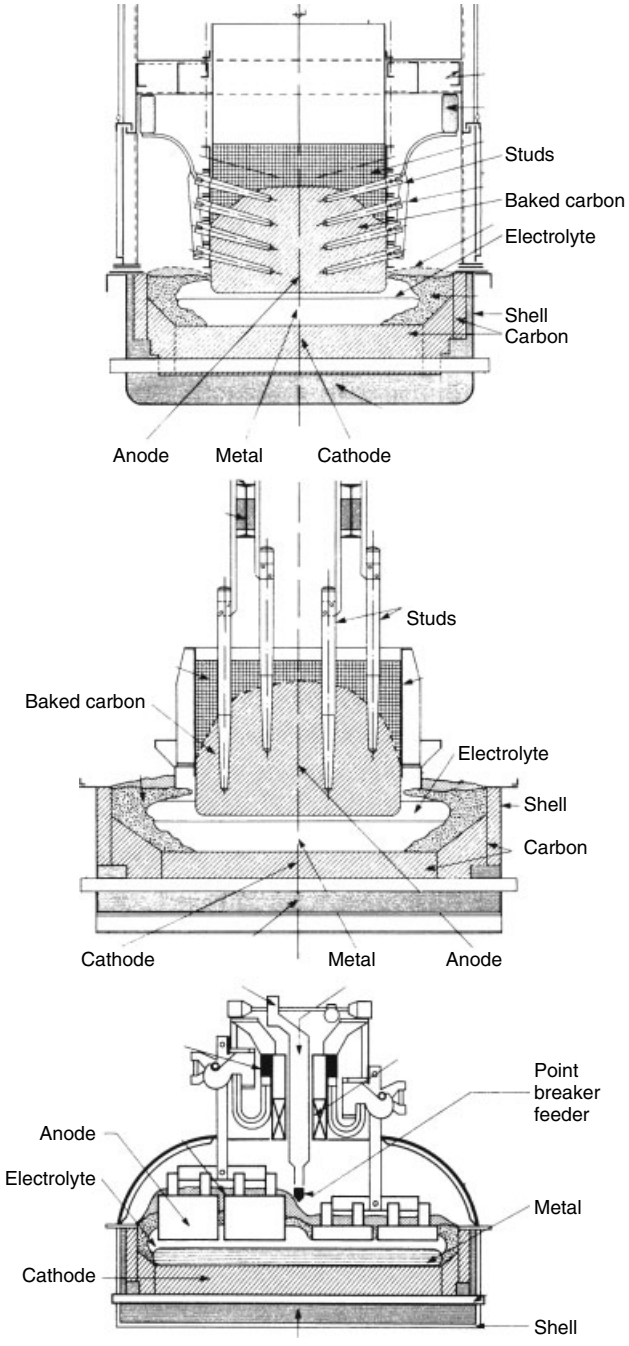


Fig. 3 Three types of Hall–Héroult cell in commercial use. The upper two use ''Soderberg'' anodes while the last uses prebaked anodes. From McGravie et al. [13].

cell shown would be 2–3 m horizontally with the other horizontal dimension being about 10–15 m. All three types use carbon anodes but the upper two employ carbon anodes (''Soderberg'' anodes) that are produced by the baking of carbon paste within the cell itself. In the third type the anodes are produced in a separate baking furnace and it is this type of cell, with ''prebaked'' anodes that is nowadays most common. The anodes are dangled into the molten fluoride electrolyte (at about 950–980 °C) from a conducting superstructure that is connected to the cathode of an adjacent cell so that the cells form a ''potline'' of perhaps 100 cells in series. Current flows down through the anodes, through the electrolyte to the layer of molten aluminum below it, into the carbon ''cathode'', through which pass steel ''collector bars'' connected to a ''ring bus'' surrounding the bottom of the cell, and thence to the adjacent cell. The term cathode is a misnomer because the true cathode is the interface between the aluminum and electrolyte where the cathodic reaction is the generation of aluminum from ions in the electrolyte. The dominant anodic reaction is the oxidation of the carbon anodes to produce carbon dioxide bubbles, which escape from beneath the anodes; the anodes are thereby consumed and are periodically moved downward and ultimately replaced with new ones (in cells using prebaked anodes). Prebaked cells have of the order of 20 anodes in two rows.

Currents within the cells are in the range of a few hundred kiloamperes and magnetic fields are generated within the cell by these internal currents and by the currents flowing in adjacent cells and busbars. All three magnetic field components are significant and vary significantly with position. They have a major effect on cell performance because of the forces that they produce by interaction with the current in the electrolyte and aluminum. These forces bring about a predominantly horizontal recirculating flow of both liquids and contribute to the deformation/instability of the interface between the two liquids. The last is rendered facile by the small difference in density between the two liquids (approximately 200 kg m^{-3}). When it is recognized that the distance between the anode and aluminum is only a few centimeters and that considerable problems (current flow with no aluminum production, anode overheating, current maldistribution) result from a short circuit, the importance of having a nearly flat, stable interface can be seen. As the anode carbon is consumed (usually irregularly) and the aluminum level in the cell rises (part of the aluminum is periodically siphoned off), the positions of the two surfaces that must be kept apart become less clearly defined, compounding the problem. Another complexity is the flow engendered by the gas bubbles released from the anode as they rise through the electrolyte. Another important phenomenon is a redissolving of a fraction of the aluminum at the aluminum–electrolyte interface and transport of the dissolved aluminum, or other reduced species, to the anodes where it is reoxidized. In this way the current efficiency (CE) (production rate as a percentage of that expected from Faraday's law) is reduced and therefore CE is a function of the electromagnetically bubble driven flow in the cell.

Yet another important aspect of flow in the cell is its influence on the transport/dissolution of alumina that is periodically added to the cell. The alumina is sparingly soluble but its concentration must be maintained at a sufficiently high

level to minimize ''anode effects'' which cause significant fluorine emission from the electrolyte. Feeding too much alumina to the cell can result in it accumulating between the aluminum and carbon ''cathode'' and interfering with the proper current flow at that point. Finally, attention is drawn to the areas at either side of the electrolyte/aluminum in Fig. 3. Here a frozen layer of electrolyte (''ledge'') protects the carbon sidewall from attack by the electrolyte. The frozen electrolyte, admixed with the alumina feed, extends over the top of the electrolyte. No practical sidewall materials have been found to resist attack by the electrolyte over the lifetime of the cell (several years) and therefore it is essential to maintain the ledge. This is achieved by a balance between the convective heat influx from the electrolyte and aluminum and a conductive outflow into the sidewall and shell and thence into the surrounding air. Consequently cell room ventilation is not only significant for the operators but also for the question of whether the ledge can be properly maintained.

It is well known that the Hall–Héroult cell is a great consumer of electrical energy, amounting to a few percent of the US national production of electricity. The energy consumption is in excess of the theoretical requirement (heating alumina and carbon to cell temperature and supplying the heat of reaction) by a factor of more than two and much effort has been expended in recent years to reduce this energy consumption. A trend that has persisted throughout the twentieth century is that of Hall–Héroult cells becoming larger. A decade or two ago, cells usually had currents of less than 200 kA; cells in the 200–300 kA range are now quite common with some cells operating at more than 300 kA. Vanvoren et al. [14] have described a 500 kA cell that appears ready for commercial use.

4.2.4 The Electrochemistry of the Hall–Héroult Cell

4.2.4.1 The Electrolyte

The electrolyte (''bath'' in the terminology of the aluminum industry) is a molten NaF/AlF_3 salt typically with an excess of AlF_3 beyond that of cryolite (Na_3AlF_6). The excess is typically of the order of 10% and it is usual to refer to the ratio of NaF to AlF_3 as the ''bath ratio'' or ''cryolite ratio'' and to specify whether this ratio is in mass or moles (although some confine the term cryolite ratio to the molar ratio). In addition, the electrolyte contains dissolved alumina as represented by Skybakmoen et al. [15], which gives the solubility of alumina in Na_3AlF_6 (Fig. 4). We see an alumina solubility of approximately 20 mol%, or 10 wt%, at 966 °C.

As discussed in Grjotheim and Welch, the dissolved alumina is in the form of oxyfluoride complexes formed by

$$F^- + 2AlF_6{}^{3-} + Al_2O_3 = 3AlOF_3{}^{2-} + AlF_4{}^-$$

and

$$3F^- + 4AlF_6{}^{3-} + Al_2O_3 = 3AlOF_5{}^{4-} + 3AlF_4{}^-$$

Furthermore the electrolyte usually contains additives (AlF_3 can be so regarded) that are intended to reduce the vapor pressure of the electrolyte, and therefore vapor losses/emissions of fluorides, or to improve such properties as electrical conductivity, liquidus temperature, or density. A representative additive is calcium fluoride. For example cryolite plus 5 wt%

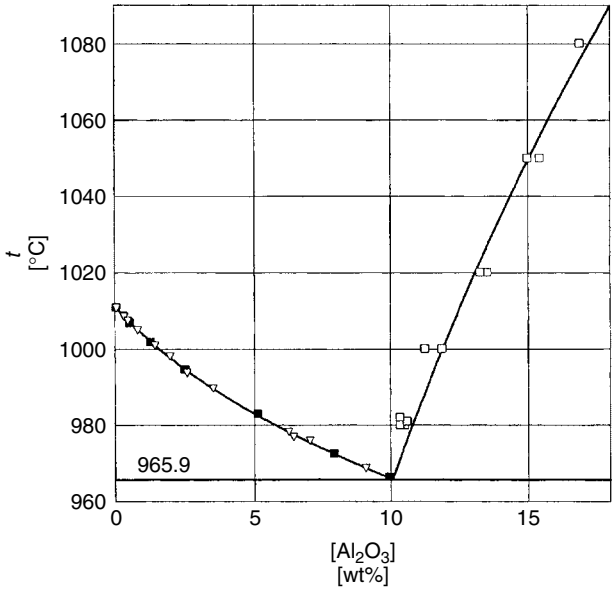


Fig. 4 Partial phase diagram giving the solubility of aluminum oxide in cryolite as a function of temperature. From the work of Sybakmoen and colleagues [15].

AlF_3 and 5 wt% CaF_2 has a minimum on its liquidus curve (at 7 wt% alumina) of about 940 °C compared to 962 °C (at 10 wt% alumina) for cryolite. Unfortunately CaF_2 additions also increase the density and lower the electrical conductivity. The former effect aggravates the tendency of the aluminum–electrolyte interface to deform, or become unstable, under electromagnetic forces and the disadvantage of the latter effect is obvious. Additions of aluminum fluoride to cryolite reduce the bath density but also reduce conductivity while lithium fluoride additions increase conductivity and slightly reduce the density. The consequence of the competing effects of these additives has resulted in disagreement on the virtues of using various additives (particularly lithium fluoride). It should also be recognized that additives could be electrolytically reduced at the cathode, to some extent, so that the product aluminum is contaminated with traces of species such as calcium or lithium that must be tolerated or removed.

4.2.4.2 The Thermodynamics and Electrode Kinetics of Hall–Héroult Cells

The thermodynamic energy requirement for reduction of alumina in the Hall–Héroult cell has been examined by several authors. Because the reaction entails an increase in entropy it is the enthalpy of reaction, ΔH_R^0, rather than the Gibbs free energy of reaction, ΔG_R^0, that determines the minimum energy consumption. Grjotheim and Welch and also Grjotheim and Kvande report that

$$\Delta H_{tot}^0 = 1.43/x + 4.91 \text{ kWh/kg Al} \quad (1)$$

where ΔH_{tot}^0 now includes the enthalpy required to take alumina and carbon at

298 K and produce aluminum, CO, and CO_2 at 1250 K. In this equation x is the CE (expressed as a fraction) so that for a CE of 100% the energy requirement is 6.34 kWh/kg Al. Equation (1) assumes that at 100% CE the anodic reaction is entirely the generation of CO_2 but that at lower CEs aluminum is consumed within the cell by the reaction

$$2Al + 3CO_2 = Al_2O_3 + 3CO$$

The "aluminum" reacting with CO_2 bubbles in this back reaction within the electrolyte may be colloidally suspended aluminum droplets, dissolved aluminum or some reduced species of aluminum present in small quantities in the electrolyte. This is not certain, but it is generally acknowledged that the reoxidation of the species by CO_2 bubbles is the principal cause of current inefficiency.

The thermodynamics of the Hall–Héroult cell have recently been reexamined by Haupin [16] and by Haupin and Kvande [17]. In the latter paper, inert anodes producing oxygen are also treated. Among the useful results presented is a correlation for the activity of alumina in the electrolyte versus the relative oxide saturation (ROS), Fig. 5.

That correlation is

$$a_{Al_2O_3} = -0.03791\text{ROS} + 2.364\text{ROS}^2 - 2.194\text{ROS}^3 + 0.8686\text{ROS}^4 \quad (2)$$

where

$$\text{ROS} = \text{Mass\% } Al_2O_3/\text{Mass\% } Al_2O_3 \text{ at saturation} \quad (3)$$

These authors report on overvoltages, including the so-called bubble overvoltage,

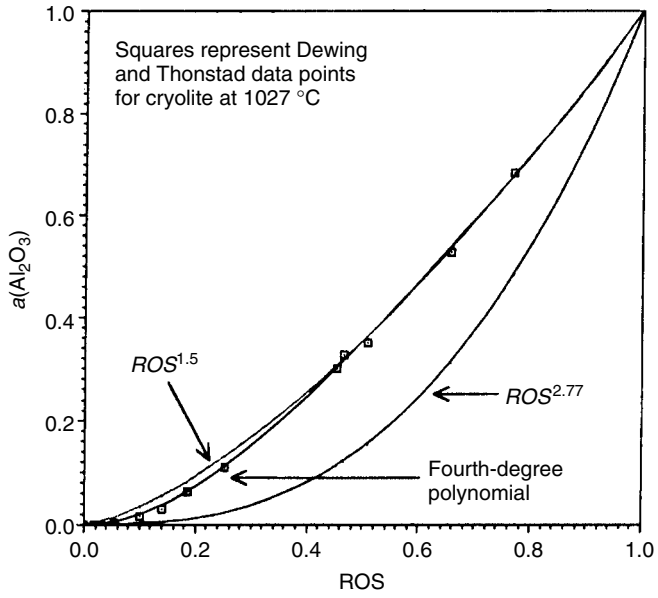


Fig. 5 Correlations (see text for equations) given by Haupin and Kvande for the activity of alumina in Hall–Héroult electrolytes [17].

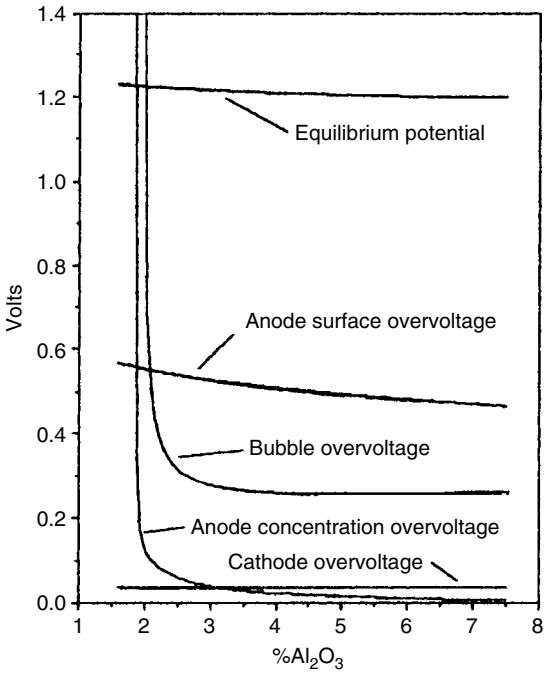


Fig. 6 Overvoltages in a Hall–Héroult cell versus electrolyte alumina content. From the work of Haupin [16].

which really represents the interference to current flow due to bubbles on the anode and within the electrolyte. Figure 6 shows some components of the cell voltage as a function of alumina content in the electrolyte for a particular electrolyte, current density, and temperature. Other components of the cell voltage appear in Fig. 7 and include the important IR drop within the electrolyte occupying the anode–cathode distance (ACD). The last is important because it is recognized that

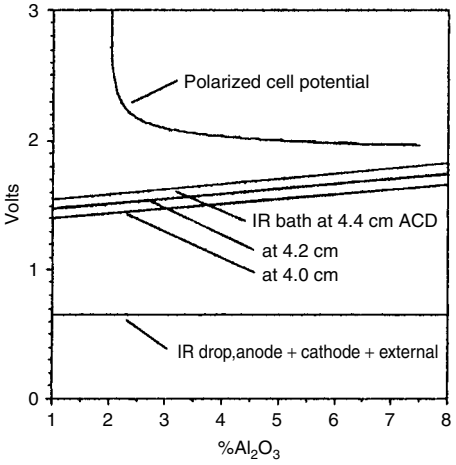


Fig. 7 Other components of cell voltage reported by Haupin [16].

one of the best opportunities for reducing energy consumption in the Hall–Héroult cell is the reduction of the ACD, therefore of the cell volts and kilowatt-hour per kilogram of aluminum. The rise of the polarized cell potential (the sum of the potentials appearing in Fig. 6) as the alumina concentration drops should be noted. It is indicative of the cell approaching a voltage at which an anode effect (discussed later) is likely to occur. This rise is frequently used in cell control to restore the alumina content by feeding the cell. Haupin and Kvande provide two figures (Figs 8 and 9), summarizing the potential drops and energy consumption in Hall–Héroult cells, including one fitted with inert anodes (Fig. 9). These calculated results pertain to a large modern cell operating at over 300 kA. The back EMF of the cell with the inert anodes (2.38 V) is large compared to the cell with carbon

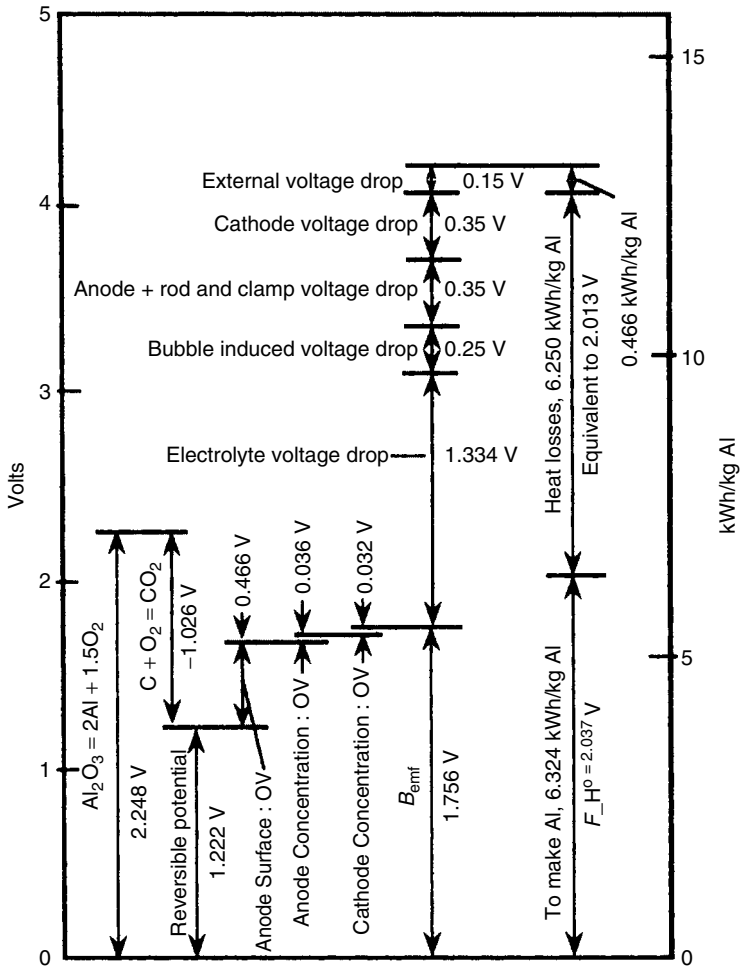


Fig. 8 Cell voltages and energy consumption for the Hall–Héroult cell, according to Haupin and Kvande [17].

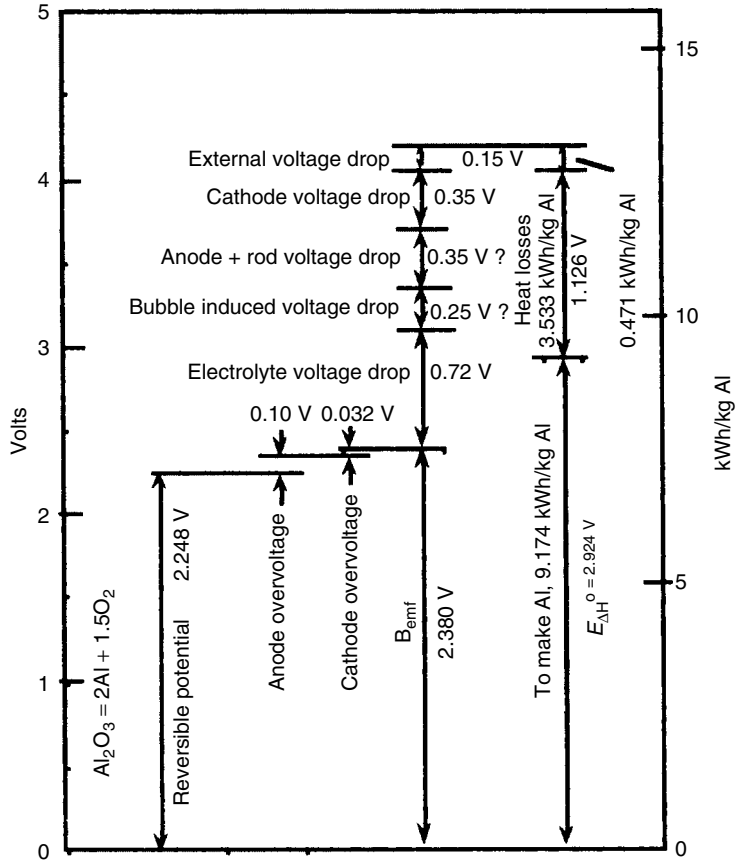


Fig. 9 Cell voltages and energy consumption for a cell equipped with an inert (oxygen evolving) anode from the work of Haupin and Kvande [17].

anodes (1.756 V), largely due to the much higher reversible potential for evolution of oxygen, compared to CO_2. This is compensated for by an (assumed) opportunity to operate at lower ACD reducing the voltage drop here from 1.334 to 0.72 V. With other voltage drops, the same for both types of anodes, the cell voltages (4.19 and 4.20 V) are virtually identical and if the CEs are the same then the kilowatt-hour per kilogram of aluminum will be the same. However, as seen in these two figures, to achieve the necessary heat balance in the cell with inert anodes the heat losses from the cell (primarily heat transport through the walls and bottom of the cell) must be reduced by half. This reduction is not easily achieved because insulation can only be added to a limited extent before the ledge protecting the sidewall of the cell is melted. Consequently, a strategy for such cells would be to increase the current, thereby increasing the productivity, but also increasing the kilowatt-hour per kilogram of aluminum. Figure 9 clearly illustrates that the success of inert anodes, in reducing energy consumption, is dependent on further reductions in the voltage drop in the ACD

than envisioned in Fig. 9 or on additional energy savings, such as reducing the 0.25 V ascribed to bubbles in this figure.

The overvoltages appearing in Fig. 8, particularly the surface overpotential at the carbon anode (low exchange current density) represent a significant contribution to the cell voltage and have been well studied as a consequence. Most investigators have represented the anodic overpotential, η, using a Tafel equation:

$$\eta = a + b \log i \qquad (4)$$

where a and b are constants and i is the anodic current density. However, as pointed out recently by Yang et al. [18], there is scatter in the published values, which was attributed to experimental difficulties. Within the range $i = 0.05–1$ A cm^{-2}, a ranges from 0.2 to 0.75 while b ranges from 0.09 to 0.6.

Modern electroanalytical techniques have been used to investigate the anodic reaction and Zhu and Sadoway [19] have provided a list.

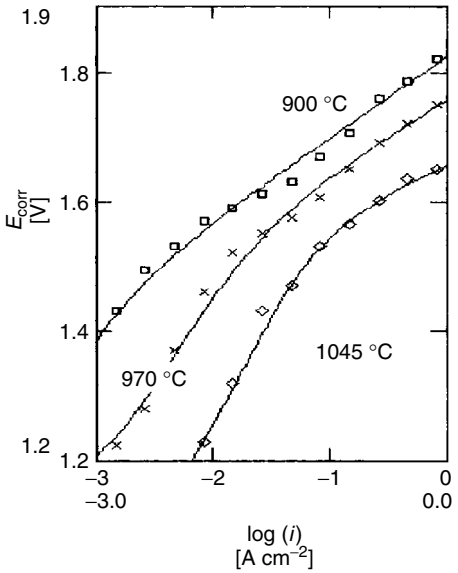


Fig. 10 Fit of the Sorensen and Kjelstrup equations (Butler–Volmer equations applied to two supposed reactions) to the anode polarization data of Solli [12].

Sorensen and Kjelstrup [20] have suggested that the anodic reaction should be considered as two reactions: a two-electron process producing CO and a four-electron process producing CO_2. The former was thought to dominate at low overpotentials while the latter would be dominant at the high current densities and overpotentials representative of the Hall–Héroult cell. Furthermore, these authors suggested that analysis of cell data should be made using the full Butler–Volmer (B–V) equation rather than just the one (anodic) term represented by the Tafel equation. Figure 10 shows the fit of their two reaction B–V equations to the experimental data of Solli [21].

A long established method for measuring the CE is the Pearson–Waddington equation [22]:

$$\mathrm{CE} = 100\% - 0.5(\%\mathrm{CO}) \qquad (5)$$

This gives the CE in terms of the %CO in the gas leaving the anodes from the stoichiometry of the back reaction

$$2\mathrm{Al} + 3\mathrm{CO_2} = \mathrm{Al_2O_3} + 3\mathrm{CO}$$

A difficulty with this determination of CE is that it is erroneous if there is significant electrochemical generation of CO at the anode; Sorensen and Kjelstrup examined the magnitude of this error based on their analysis of the two-electron generation of CO.

An alternative to CE measurement by measuring the CO/CO_2 in the anode gas is the measurement of the total oxygen content of the anode gas. While probably impossible on an industrial cell

because of air diluting the anode gas as it leaves the cell, this can be achieved in closed laboratory cells by, for example, mass spectrometry of the exit gas. This CE measurement is unaffected by the electrochemical production of CO, or by any Boudouard reaction

$$C + CO_2 = 2CO$$

Dorreen and coworkers [23] have used this oxygen balance method to measure CEs in laboratory cells where the alumina content of the electrolyte was allowed to fall to the onset of an anode effect. Simultaneously they measured CE by the Pearson–Waddington approach and found that this latter method yielded a significantly lower CE. By the oxygen balance method the CE was found to be unaffected by alumina content, or by ACD, but to show a strong increase with increasing current density and a strong decrease with increasing cryolite ratio (increasing NaF to AlF_3 in the electrolyte).

Kisza and colleagues [24] have recently reported on the kinetics of both the anodic and cathodic reactions. They used both the impedance spectroscopy and the galvanostatic relaxation method in their experiments. Figure 11 is their plot of the dependence of the exchange current density for the cathodic reaction on the alumina content of the electrolyte. They used their results to suggest mechanisms for cathodic and anodic reactions.

4.2.4.3 Anode Effects and Gas Emissions

Hall–Héroult cells are usually operated with intermittent feed of alumina. The consequence is that, except at times when feeding is occurring, the alumina content of the electrolyte is dropping. If the cell is not properly fed the alumina content drops to a point where an ''anode effect'' occurs. Effectively the resistance of the cell increases; because the cell is just one of many in the series, the consequence is that the cell current falls by only a minor amount when the anode effect occurs but the cell voltage increases several fold. Under this condition the cell is consuming much more energy than is normal and the heat generated in the cell is excessive, running the danger of melting the ledge and attacking the sidewall, if the effect persists for more than a few minutes. Finally, as discussed here, it is during anode effects

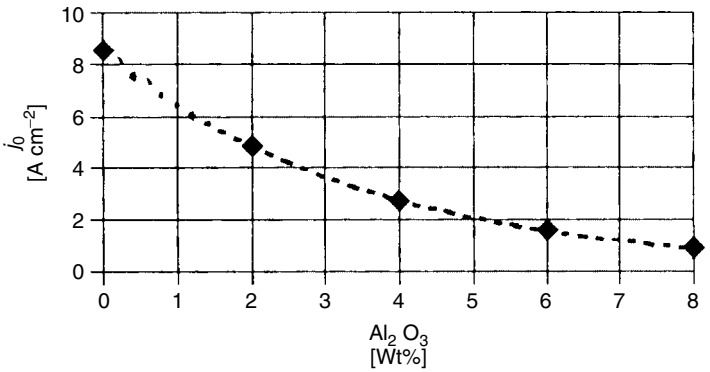


Fig. 11 Dependence of exchange current density for the anodic reaction on electrolyte alumina content as reported by Kisza et al. [24].

that the most serious emissions occur. Anode effects are quenched by adding alumina usually with modest agitation, such as raising and lowering the anodes, to bring dissolving alumina to the anode surfaces. The alumina content at which an anode effect occurs is dependent on anode current density but is typically in the range 0.5–2.2 wt% [7]. In the past, anode effects were allowed to occur, despite their disadvantages, mostly because they provided a crude measure of the alumina content of the electrolyte, permitting the operator to adjust the rate of alumina feeding and ensuring that overfeeding of alumina would not occur. Furthermore, sparking and high temperatures occurring at the anode surface during an anode effect would remove unevenness on the bottoms of the anodes and the agitation of the cell contents would bring carbon particles (from anodes) to the surface where they would be removed by oxidation. Nowadays, automatic control of electrolyte alumina content, based mostly on tracking cell voltage, has advanced to the point where anode effects are less necessary and there are efforts by the industry to decrease both their frequency and their duration.

The incentive for minimization of anode effects is the reduction of the emissions of the perfluorocarbon compounds (PFCs) CF_4 and C_2F_6, which are known to be greenhouse gasses. Although Marks et al. [25] report low levels of such emissions during normal cell operations (0.1 to 0.3 ppm CF_4 in the cell exhaust ducts, compared to up to 870 ppm CF_4 and 137 ppm C_2F_6 during anode effects), it is during anode effects that these emissions are most extreme. Tabereaux [26] has estimated global aluminum industry emissions of CF_4 and C_2F_6 as 30 000 and 3000 tonnes year^{-1}, equivalent to the global warming potential of 223 million tonnes of CO_2. Thonstad and coworkers, in a recent review of anode effects [27], contemplated the effect of proposed "CO_2" taxes at \$15/tonne and estimated an increased production cost of 1.2% per anode effect per cell per day.

In their survey of emissions from six aluminum plants, Marks et al. included a range of technologies (cells with vertical stud Soderberg anodes, horizontal stud Soderbergs and prebaked anodes with the last including cells equipped with both the older feeding of alumina at the sides of the cell and cells with modern central point feeders). Cell currents ranged from 68 to 322 kA. The CF_4 emissions were found to correlate well with the minutes per cell per day that anode effects were occurring. The results suggested that the reduction of both the frequency and duration of anode effects would reduce emissions of PFCs but the former strategy would be more effective because emission rates are high early in the anode effect. The success of one company in achieving this goal during the first half of the nineties is shown in Fig. 12, taken from the paper by Marks [28]. Recent reports on PFC emissions are those of Marks and colleagues [29] and of Dolin et al. [30].

There have been numerous fundamental studies of the anode effect and the consensus is that the anode is separated from the electrolyte by a resistive film when it is at the high voltage of the effect. Experiments with see through cells show that, as the alumina content of the electrolyte drops, the carbon anode becomes less wetted by the electrolyte. Eventually a gas film extends over the whole anode under these circumstances. However, it is possible that this gas film is not the cause of the effect because artificially bubbling additional gas under an anode does not bring about an anode effect. The transition

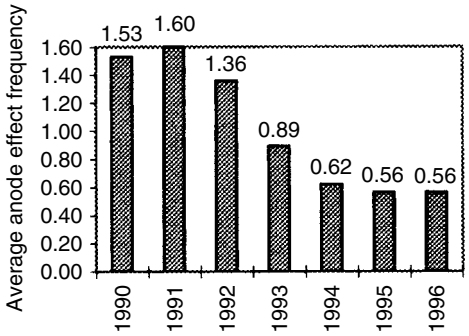


Fig. 12 Frequency of anode effects (anode effects per cell per day) at Alcoa versus year. From the paper of Marks [28].

from normal operation to anode effect has been examined by Zhu and Sadoway [31] using a variety of electrochemical techniques, along with the analysis of cell gasses emanating from their laboratory cell. Figure 13 contains some of their results. These are from experiments where the anode current was measured as the anode was swept, starting from rest potential, toward more anodic potentials (with

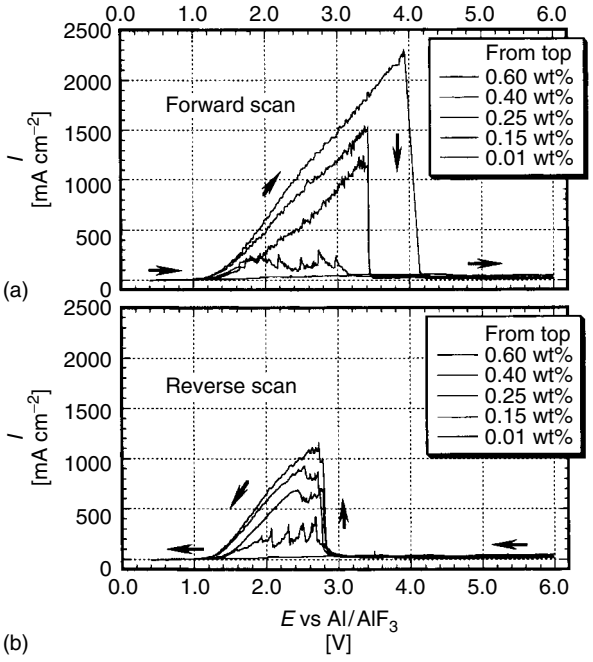


Fig. 13 Linear potential sweep results of Zhu and Sadoway showing a large drop in current at an anode as the potential reaches a few volts (versus Al/AlF_3) during the forward sweep and the persistence of the low current as the potential is subsequently swept downward. Note the dependence on alumina concentration [31].

respect to an Al/AlF_3 reference electrode), see 13(a), and then, at 6 V, the sweep was reversed down to the rest potential. The experiments were carried out at three different alumina levels. The significant result is that, on the forward sweep, the current plummets once a certain (alumina concentration dependent) potential is reached. Prior to that critical potential the anode is undergoing the normal electrochemical generation of CO_2 and CO. Beyond that potential, the anode is virtually passivated, at least up to 6 V. That passivity is sustained as the anode potential is lowered on the reverse sweep until lower potential than would sustain current on the upward sweep. The potential at which the passivity ends appears to be independent of alumina content.

Assuming the applicability of these results to a real cell the following informal stability analysis of the present authors becomes possible. Take a cell operating at 1000 mA cm^{-2} (which would be representative of an industrial anode) in an electrolyte containing 0.6 wt% alumina. Figure 13(a) indicates that the anode could be operating in one of two possible states: at a potential of 2.4 V or at a potential slightly over 4 V. Now consider a perturbation upward in anode potential, as might occur if the anode, or one of it neighbors, becomes a little more covered in gas. At the 2.4 V state, this perturbation can be accommodated; the current density on the still active area rises to compensate for the drop in current elsewhere and the cell reaches a new, slightly different, steady state. At the 4 V state things are very different. A perturbation upward in voltage now *diminishes* the current on the active anode area. This then causes a further increase in potential (as the potline seeks to maintain its current), a further current reduction on the active area, leading to further potential increase and so on until the potential has progressed to the right-hand side of this figure, thus passivating all active regions.

More results of Zhu and Sadoway appear in Fig. 14 where the cell current with the anode potential held at a fixed value (rather than swept) is plotted against that potential, along with the composition of the exit gasses. The results extend into the high voltages of the anode effect and are for a low level of Al_2O_3 in the electrolyte. Up to about 2.5 V the normal generation of CO_2 and CO (the latter dominant in this experiment because of the Boudouard reaction) are seen. Beyond that point CF_4 starts to appear but, in conformity with the work of others, C_2F_6 is low (below the detection limit here). From roughly 3.5 to 4 V, the current is lower with increasing potential, reducing the generation of all three gasses. Beyond 5 V, there is some recovery of the current and of CO generation but the significant effect is a great increase in CF_4 generation.

A very recent fundamental investigation of the anode effect is that of Vogt [32] who showed that the effect of alumina concentration on the critical current density (above which an anode effect occurs) can be explained by the alumina mass transfer, fluid dynamics of gas release and wettability of the anode.

The upturn in voltage of an industrial cell that marks the start of an anode effect is seen in Fig. 15 from the paper by Dorreen et al. [33]. If the horizontal axis were linear, rather than logarithmic, then a sharp increase in voltage, starting a few tens of seconds before the anode effect, would be apparent. These authors point out that emissions from Hall–Héroult cells were investigated as early as the 1950s with Henry and Holliday [34] reporting emissions of PFCs only during anode

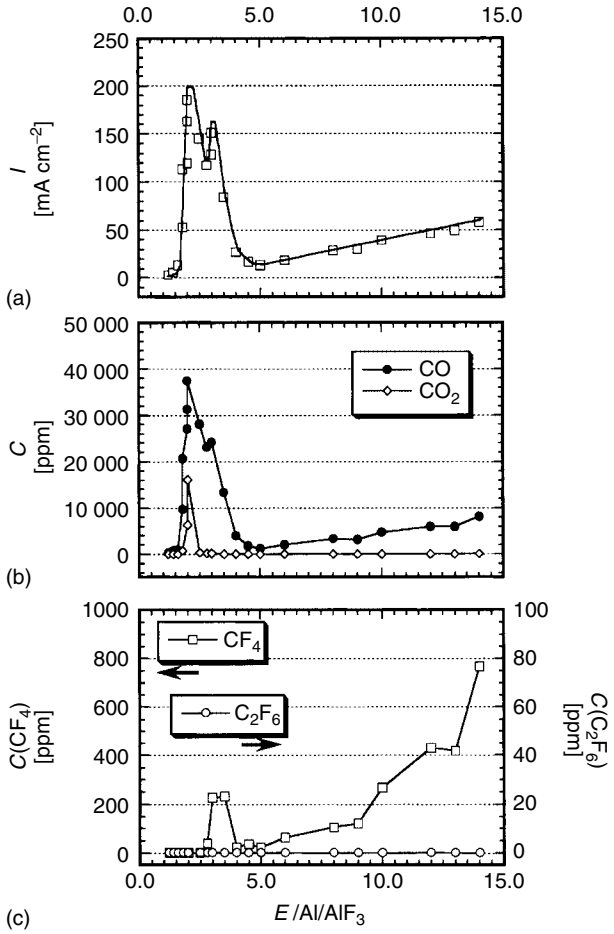


Fig. 14 Potentiostatic results of Zhu and Sadoway with cell current and off gas composition plotted versus (fixed) cell potential [31].

effects. Henry and Holliday also investigated emissions of sulfur-containing gasses, such as carbonyl sulfide (COS) (now thought to be detrimental to the ozone layer) which they detected in unburned gasses during normal operation. Dorreen and colleagues provided a list of possible anodic reactions with standard electrode potentials at 970 °C for those that were electrochemical, and equilibrium constants for those that were not. They also carried out experiments to monitor the emissions of PFCs and sulfur-containing gasses from a laboratory cell using mass spectrometry. The cell was provided with an argon purge to prevent oxidation of gas products by air. Figure 16 is representative of the results of these authors and shows emissions before and during an anode effect. Numbers on the curves are the mass per electron in the mass spectrometer and in some cases (e.g. 28) this leads

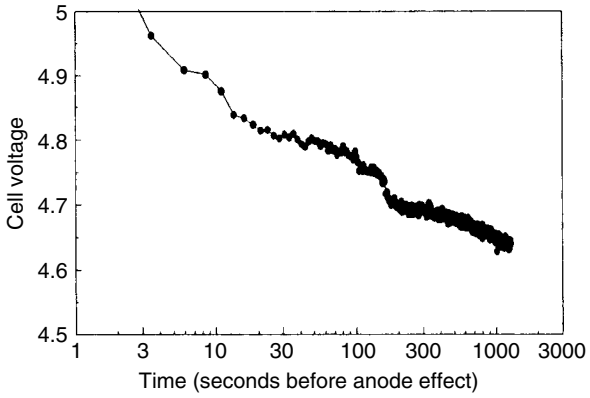


Fig. 15 Increase in cell voltage before an anode effect as reported by Dorreen et al. [33].

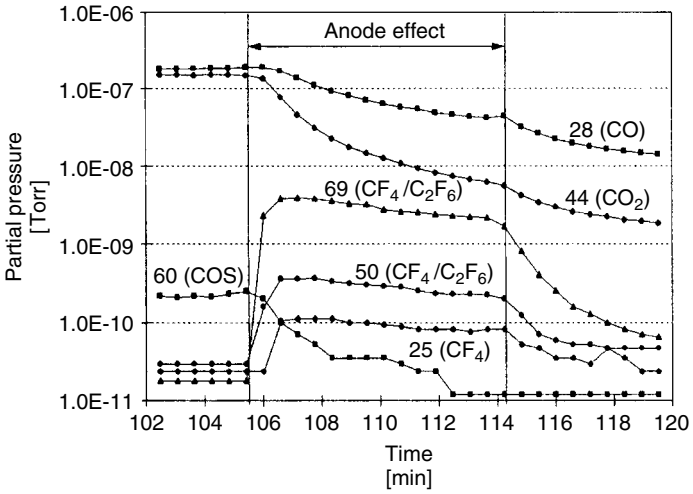


Fig. 16 Gasses emitted by a laboratory cell before, during and after an anode effect, as measured by Dorreen and colleagues using mass spectrometry [33].

to an unambiguous identification of the emitted species (e.g. CO). In other cases (e.g. 35) some uncertainty exists concerning the origin of the ionized species (e.g. CF_4 or C_2F_6) in the mass spectrometer. The logarithmic scale of the vertical axis should be noted and the fact that, in normal operation, CO_2 and CO (the latter enhanced by the Boudouard reaction) are the principal emissions, with small amounts of carbonyl sulfide and amounts of PFCs only at the detection limit. Emissions of all but PFCs drop during the anode effect because of a drop in cell current as the voltage rises (the power supply could not sustain constant current at high voltages). PFC emissions rise sharply during the anode effect as found in other investigations such

as those discussed earlier. One notable result of this investigation was the detection of COF_2 starting a few minutes *before* the anode effect and continuing throughout the effect.

Utne and coworkers [36] measured the emissions of carbonyl sulfide from industrial cells during normal operation. High levels of COS were found in the undiluted/unoxidized anode gas; amounts corresponded to 3.3–10.2 kg tonne^{-1} of aluminum produced. Most of the sulfur emitted from the cells appears to originate from sulfur in the anodes. However, the COS is oxidized as the anode gasses leave the cell. After the gasses pass through dry and then wet scrubbing, the COS content corresponds to 0.21–0.26 kg tonne^{-1} Al.

Hyland et al. [37] have recently reviewed sulfur and fluoride emissions (including particulate emissions) from Hall–Héroult cells and concluded that operational changes over the past few years, such as a tendency toward lower ratio (more volatile) electrolyte may have made emission control more difficult.

4.2.4.4 Gas Bubbles in Hall–Héroult Cells

Many effects of gas bubbles released at electrodes (on electrolyte flow, mass and heat transport, conduction, etc.) have been well studied in the past. A text with an extensive treatment of this topic is that of Hine [38]. However, in Hall–Héroult cells these effects are worthy of special mention because the relatively high current density, of the order of 1 A cm^{-2}, and temperature make the volumetric gas evolution rate from the anode large. Furthermore, difficulties of measurement on actual cells mean less knowledge of these effects than in many other electrochemical cells. Finally, one effect of the bubble is to make the task difficult in reducing the enormous energy consumption of the Hall–Héroult cell.

Carbon dioxide bubbles nucleate on the underside of the carbon anodes (and to a lesser extent on the sides), grow through a near spherical shape and start to coalesce to form large bubbles that are more disc-shaped than spherical. These bubbles, now occupying a significant fraction of the anode area, then sweep across the anode underside to an edge where they escape. As they do so, they accumulate more gas from smaller stationary or slower bubbles in their path and leave behind bubble free regions where fresh nucleation can occur. These dynamics are well described by Fortin et al. [39].

Any bubbles in the space between the anode and the cathode will interfere with current flow and effectively raise the cell voltage. This effect has been studied by several investigators in the past but one of the investigations most relevant to the Hall–Héroult cell is that of Hyde and Welch [40] who placed ceramic spheres and discs in a cell where the electrolyte was a lead chloride–sodium chloride eutectic. Current was passed between a liquid lead pool (simulating the aluminum pool of the Hall–Héroult cell) and a horizontal carbon cathode placed in the electrolyte above it. At this electrode, passage of current results in deposition of lead. The ceramic shapes floated into the upper part of the space between the electrodes and were therefore positioned much as we would expect the bubbles to be beneath an anode. The current interruption method was used to measure the cell resistance. Figure 17 shows representative results where cell resistance is plotted against the ACD for the cell without ceramic ''bubbles'' (Bath only), with 3 large spheres, with 6 large spheres, and with 6 large spheres plus 44 small spheres. The parallel curves reflect the fact

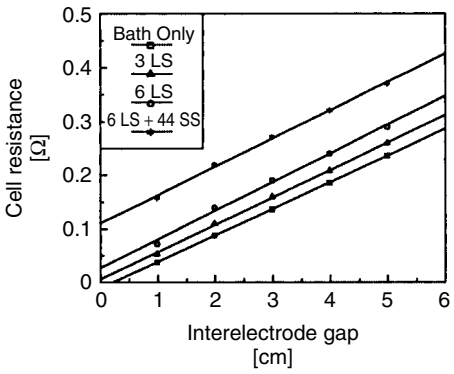


Fig. 17 Cell resistances measured by Hyde and Welch on a model cell. Resistance is reported as a function of interelectrode gap with three large ceramic spheres (simulating gas bubbles) in the electrolyte, 6 large spheres, 6 large and 44 small spheres, or no spheres [40].

that the spheres introduce a fixed additional resistance into the interelectrode gap independently of the size of that gap. These investigators found that this additional resistance was proportional to the volume of the "bubbles" and nearly independent of their size and shape.

From the description of bubble growth and release in the penultimate paragraph one would expect individual bubbles to be detectable, for example, in the cell voltage trace for small anodes, such as those used in laboratory cells. Aaberg and colleagues [41] used a cell with a 10-cm diameter graphite anode and measured both cell voltage and electrolyte level as a function of time. Both these variables showed periodic fluctuations and the latter fluctuation enabled the bubble volume at release to be measured. Bubble release frequency was found to increase with increasing current density, ranging from 0.13 to 3.2 Hz. Bubble volume at release also increased with current density but only slightly. Simultaneous measurement of the cell resistance by current interruption yielded resistances for the anodic bubble layer of 0.5–1.5 Ωcm^2.

Wang and Tabereaux [42] too used cell voltage fluctuations (for a 15-cm diameter anode) to study bubble release. Figure 18 shows the voltage trace for an anode after 1200 Ah of electrolysis (upper plot) and for an anode after 12 460 Ah at the same current. A bubble release frequency is clearly discernible in Fig. 18(a), while frequencies are distributed in the lower trace. These investigators found an increase of bubble frequency and of bubble volume with current density for much of their data, but a minimum in bubble frequency with flat anodes (little rounded by electrolysis) between 0.7 and 0.8 A cm^{-2}.

Xue and Øye [43] have described an investigation of bubbles on industrial cells by analysis of the sound emitted from anodes.

As described in a section later, there has been a great interest in conserving the electrical energy entailed in aluminum production by reducing the energy expended in merely passing the current through the electrolyte between anode and cathode. This IR drop (in voltage) accounts for more than 30% of the total energy consumption in a modern cell (see Fig. 8) and an even higher percentage in older cells. An obvious strategy is to reduce the ACD but the results of Fig. 17, and similar results from previous studies (e.g. Shekhar and Evans [44–46]) suggest that pursuing this strategy may only bring limited gains. The interelectrode resistance is *not* proportional to the ACD so that halving the ACD

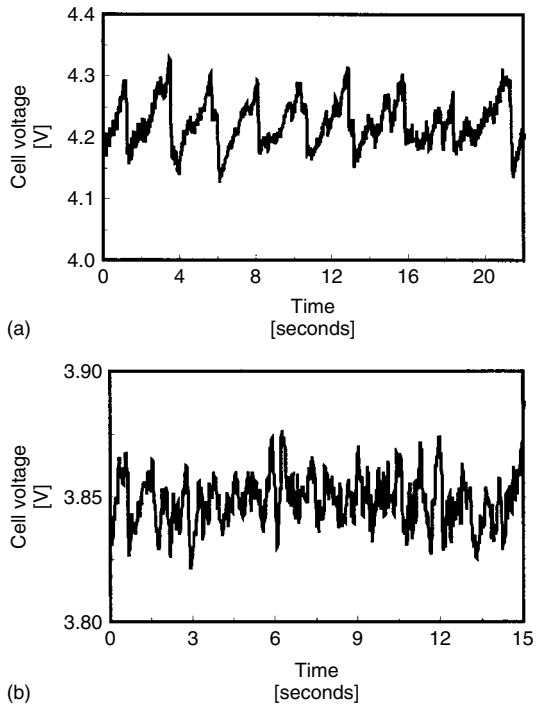


Fig. 18 Voltage traces for a new anode (a) and an older (15-cm diameter) anode (b) as reported by Wang and Tabereaux. The release of individual bubbles can be seen in the former trace [42].

does not halve the resistance as might at first have been supposed. Figure 19, from the pilot scale studies of a cell with novel electrodes by Dorward and Payne [47] illustrates this point. The plot is of the measured interpolar resistance presented by the electrolyte versus ACD. In a bubble free electrolyte, the interpolar resistance should be proportional to ACD (ignoring minor fringing effects). With anode

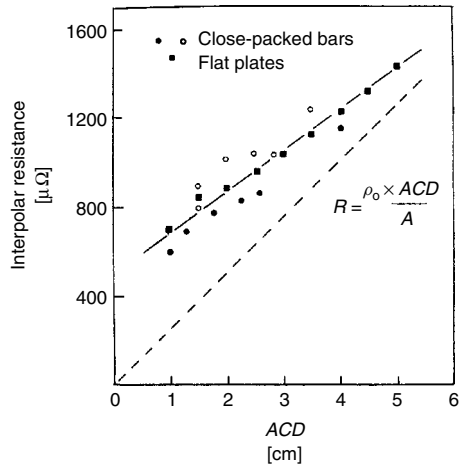


Fig. 19 Plot of interpolar resistance versus anode–cathode distance from the work of Dorward and Payne on a pilot cell equipped with "wettable" cathodes. The difference between the broken and solid lines is due to gas bubbles [47].

gas bubbles present, the bubbles occupy a greater and greater fraction of the inter-electrode gap as the ACD is reduced; the apparent resistivity increases so that the interpolar resistance is greater than the proportionality represented by the broken line, at low ACD. The behavior of bubbles therefore has a direct impact on energy conservation efforts and modeling of this bubble behavior is described in a section below.

4.2.5 Mathematical and Physical Modeling of Hall–Héroult Cells

4.2.5.1 Steady-state Models for Fields, Flows, and the Interface

The multiphysics and multiscale character of the important features of Hall–Héroult cell operation makes difficult laboratory scale experimentation that is relevant to industrial pot operations. For example, cell CE is influenced by the cell-scale flow of the metal and electrolyte, which is determined in turn by the magnetic field which depends on the entire cell current. CE also depends on the finer scale flow due to release of the carbon dioxide bubbles from the anodes. It is generally not possible to examine these two effects simultaneously in the laboratory. Also, the generally hostile environment inside Hall–Héroult cells makes experimentation difficult, and the high cost of modification of full-scale pots further complicates industrial trials. In this environment, numerical or mathematical modeling of pots would be expected to be a useful tool.

On the other hand, the multiphysics, multiscale character of the problem poses formidable challenges for numerical simulation also. The consequence is that practical models are based on judicious decoupling of some parts of the physics. The importance of these uncouplings can be checked and rectified to some extent *a posteriori*. Typically, models are built in a modular fashion, with a sequential flow of information from one to the next. In addition, although the cell operation is inherently dynamic, models are often steady state, with dynamic considerations either as a follow-on or as output from simplified models. In a typical sequence, the first step is often a thermoelectric model that gives a map of the current distribution. Here, the shape of the frozen ledge is important in determining possible horizontal currents in the metal pad. An important decoupling is the effect of convection on the shape of the ledge. Another decoupling is to ignore the effect of the fluid flow in inducing current in the fluid layers; this effect is often added back after the velocity is calculated, but the effects of the induced currents are not then included in the nonlinear ferromagnetic effects in the magnetic field calculation, which is the next module in the sequence. As a next step, the flow, including the metal–bath interface deformation, is calculated. As a final step, some assessment of the magnetohydrodynamic (MHD) stability is made.

Historically, the mathematical modeling of the Hall–Héroult cell became feasible in the 1970s when computers of sufficient power to examine many of the important phenomena were first available. In spite of a great deal of progress and much development effort, some critical issues remain in relating the output of the models to operational quantities of interest, such as noise and CE. In the following sections, each of the areas is discussed in more detail.

Historically, current density distributions were usually calculated in two dimensions (vertical and the short horizontal dimension) by numerical solution of Laplace's equation for the electrical

potential between the anode surface and the ends of the collector bars; see, for example, Tarapore [48] or Evans et al. [49]. The user would specify a side-ledge profile. Presently, there is more emphasis on thermoelectric modeling to include effects of operating parameter changes such as metal pad depth on the ledge profile. The current in the buswork, sometimes including adjacent cells would be specified. In this way, the features of busbar design and cell placement could be incorporated into the computations (Lympany and Evans [50], Potocnik and Evans [51]). Presently, the calculation of the currents in the full 3D buswork simultaneously with the cell interior has become feasible.

In addition to determining the current distribution for magnetic and flow modeling, thermoelectric models of Hall–Héroult cells have as objectives side ledge and isotherm prediction to help assure cathode life. Use of such models for component design is illustrated for a collector bar/cathode block assembly by Bos and coworkers [52]. An expanding use of these models is to examine stress in potshells, where the thermal load can be augmented by a representation of the swelling of the cathode induced by sodium intercalation. Trends in this area include more simultaneous physical phenomena and more complicated geometry, such as entire quarter pots, as exemplified for a 400 kA cell in a recent paper by Dupuis [53], and moving the modeling capability closer to the plant floor as shown for the Russian industry by Sherbinin and colleagues [54] who also described a 3D thermoelectric model.

The next step in the sequence is the calculation of the magnetic field. This remains a computer-intensive task. Previously, the Biot–Savart law was used to calculate the magnetic fields within the cell. A difficulty arising at this point is how to allow for the influence of the ferromagnetic steel shell on the magnetic field. The earlier, rather crude solution to this difficulty was to modify the contributions, to the field within the cell, from external currents by a ''shielding factor'' chosen to give reasonable fits to experimental measurements. More recently, a number of approaches have been used to accommodate the steel. Ziegler and Kozarek [55] reported use of a magnetic vector potential method executed in commercial software to calculate the effects of the nonlinear magnetization of the steel. This technique suffers the drawback of requiring the meshing of the air surrounding the cell, which increases both the complexity of the meshing task and the computational requirements. Another technique that has the same problem is the two-scalar potential method, used for Hall cells by Guohua and Dexiang [56]. An approach that requires meshing of only the steel parts is a volume integral method and has been applied to Hall cells by Segatz and Vogelsang [57] and Boivin [58]. Variations on this method have also been reported by Kalimov and colleagues [59] and by Sun Yang et al. [60]. In spite of these computational advances, an ongoing difficulty of this calculation is the hysteresis of the magnetization of the steel. Therefore, the influence of the steel is not just a function of the present magnetic fields to which the steel is exposed, but also past fields. Furthermore, the permeability of the steel is temperature dependent, representing an additional coupling of the various solutions.

In this area, purpose built programs still seem to predominate, in comparison to the use of commercial software, which is more in evidence in the thermoelectric and flow calculations.

The next step in the sequence, the flow calculation, generally starts with the Lorentz forces. These are given by the cross product of the current density and the magnetic flux density, and are input as momentum source terms into the fluid flow equations. These are frequently time-averaged versions of the Navier–Stokes equations with turbulence accounted for by $k-\varepsilon$, or similar, models. Finally, the pressure fields for both the electrolyte and aluminum having emerged from the fluid flow calculations, the deformation of the interface between them can be obtained. Usually, in the interface calculation, only the larger length scales are resolved, so that the interface is only slightly curved and surface tension effects can be neglected. This calculation has not changed much – an early example modeled by Lympany and Evans [50] is shown in Fig. 20, which shows the busbar arrangement around six nearby cells. Each of the cells contains 18 anodes (the vertical anode rods from which the anodes are suspended in two rows can be seen in the figure). These are "end riser" cells where the "riser", bringing current upward from the level of the ring bus of the adjacent cell, to the anode superstructure of a cell, makes its connection at the end of the superstructure. The computed electrolyte–metal interface is depicted, for the cell indicated by the arrow of Fig. 20, in Fig. 21 where the upward bowing of the interface near the center of the cell, with a trough-to-peak distance of 109 mm, is seen. As a final step in this calculation, the mass transport of dissolved aluminum and thereby CE was calculated using a relationship based on the turbulent kinetic energy of the flow. The computed CE for this design was 84.4%, which is within the range of expectations. However, effects specific to Hall cells, such as the bubble release, which are expected to be important in determining the CE are not accounted for.

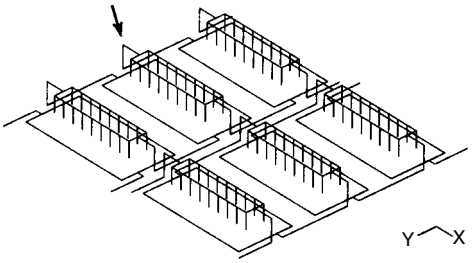


Fig. 20 The arrangement of busbars and anode rods for six "end riser" cells mathematically modeled by Lympany and Evans. The cells are in two "potlines" and each cell has 18 anodes [50].

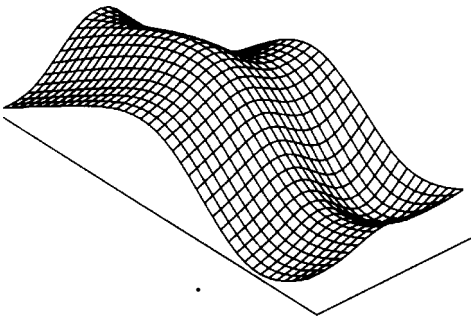


Fig. 21 The deformation of the metal–electrolyte interface for one of the cells of Fig. 20 as computed by Lympany and Evans. The trough-to-peak distance is 109 mm [50].

A more recent example is by Potocnik and Laroche [61], which addresses calculation of the steady velocity and its comparison with measurements. The authors remark on the limitation to the accuracy of these models imposed by uncertainty in the current distribution.

As in the case of the current distribution calculation, the trend in this area is the inclusion of more comprehensive physics, principally the bubble driven flow effects. A relatively early example of this was published by Fraser et al. [62], and a recent example is by Bech et al. [63]. Although inclusion of this effect was first reported in 1990, it is still computationally daunting for production calculations. In addition, serious questions remain about appropriate constitutive relationships for the bubble containing flow, particularly beneath the anodes. Also finding more inclusion are effects that depend on the velocity field; convective transport of heat and dissolved/undissolved alumina.

The rapid dissolution of the alumina fed to a cell is important for proper cell operation. However, it is generally accepted that at least a portion of the alumina fed finds its way to the bottom of the cell and forms part of a moving ''sludge'' before dissolving. The movement of the metal transports the sludge. The behavior of alumina has been examined by Kalgraf and Tørklep [64], who applied the theory of sediment transport to the alumina and performed experiments where they simulated alumina particles in the metal by nylon particles placed on a moving platform in water. Figure 22 is representative of their results with the sediment transport velocity a function of the relative velocity between the water and the platform. These same investigators and Nordbo [65] went on to compute alumina distributions in Hall cells based on computations of the electrolyte and metal flow in the cell. This paper also has an account of radiotracer measurements conducted on a Hall cell to measure velocities and turbulent dispersion of the tracer.

The works described above are based on solution of the 3D flow equations, often with additional partial differential equations for the turbulence model. These models are often executed with commercial software. An alternative to the computational treatment of the turbulent flow of the electrolyte and aluminum in the cell is the ''shallow water'' approximation described by Moreau and Evans [66] and later extended by Moreau and Ziegler [67]. In this approximation, the flow in each layer is treated by averaging the Navier–Stokes equations over the thickness of the layer (this vertical dimension of the order of 20 cm is small compared to the meters of the horizontal dimensions). Then the usual laminar/turbulent viscous terms in the Navier–Stokes equations are replaced by drag terms proportional to the velocity in each layer representing friction with its bounding surface – either the anode bottoms or the cathode top – and proportional to the velocity difference between the two layers. In this way the coupling between flow in the two layers is achieved. For simplified geometries, analytical results, or ones entailing relatively simple computations, are possible. Moreau and Evans were able to show that the channels between anodes are important conduits for electromagnetically driven flow. In this way they change the pressure distribution in the electrolyte and thereby the interface deformation. An important advantage of this approach to calculating the flow is the major increase in computational speed that it makes possible. This can especially be exploited in transient simulations as described below.

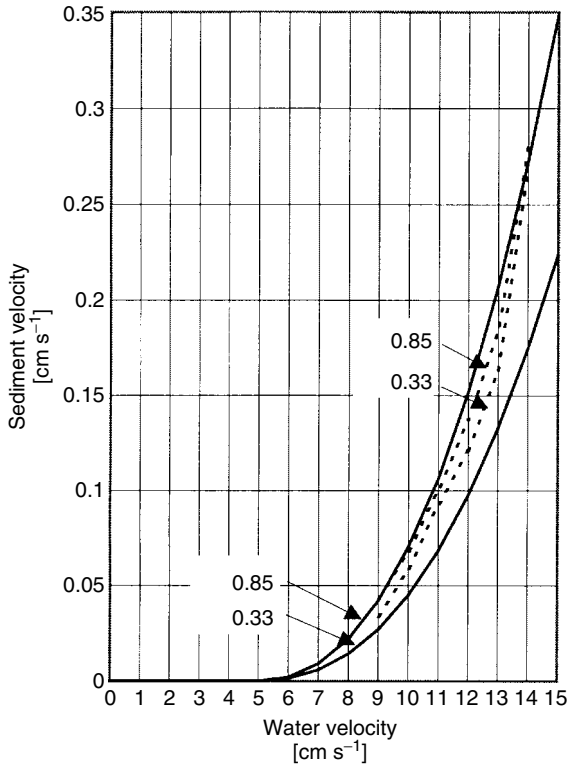


Fig. 22 Results of an investigation, by water modeling, of Kalgraf and Tørklep, into transport of alumina particles (sediment) along the bottom of a cell. The parameters on the curves are the diameters (millimeters) of nylon particles (simulating alumina particles). Solid lines are calculated curves [64].

4.2.5.2 **Models for Cell Instability**

As described above, instability of the interface between the electrolyte and molten metal is a significant problem that is one root cause of the energy inefficiency of Hall cells. Expressed simply, the interface is deformed by the electromagnetic body forces arising from the interaction between currents in the cell and the magnetic field. The currents are themselves affected by the interface position because it determines the distance between the top surface of the aluminum and the bottom of the anode. There is therefore the possibility that interface deformation leads to further interface deformation. Other mechanisms for generating waves at the interface may be significant, for example, the Kelvin–Helmholtz mechanism which stems from a shear at the interface.

The importance of the interface instability has led to many models being developed to examine the deformation of the interface and the evolution of that interface. Most such models can be placed into one of the two categories: (1) fundamental studies of the instability of the interface based on idealization of the cell and (2) computationally intensive investigations where models akin to those of the last section, incorporating most features and phenomena of the cell, are extended to time dependent calculations.

In using the former approach, it is necessary to make sufficient idealizations of the cell to arrive at a problem that is simple enough to solve. Consequently,

various authors, having made different simplifying assumptions, have found a range of instability mechanisms that may be operative in industrial cells. For example, in the early work by Sele [68], the magnetic state was represented as a uniform vertical field. He represented the interface as a tilted surface that was rotating about a vertical axis, that is, an interfacial wave traveling in the azimuthal direction and determined that such a wave could be driven by the interaction of horizontal current components and the vertical magnetic field. In this case, the horizontal current components arise from the transient deformation of the interface. In other early efforts along these lines, Urata [69] carried out a stability analysis for the interface that led to the conclusion that horizontal gradients in the vertical magnetic field were destabilizing while Moreau and Ziegler [35] examined the effect of the diversion of vertical currents into the horizontal direction on interface stability. Sneyd [70, 71] has worked on the stability of the interface in Hall cells examining instabilities that arise from both horizontal and vertical gradients in the magnetic field.

All of these causes may be expected in practical Hall cells. The vertical magnetic field is caused by horizontal currents around the cell (adjacent potline, ring bus, etc.) and by position dependent horizontal current in the aluminum and collector bars; the field varies with proximity to these currents. The current flow is predominantly vertical across the electrolyte between anode and metal, because of the high resistivity of the electrolyte, but develops significant horizontal components on entering the much more conductive metal. In practice, all of these mechanisms can be related to instabilities and their correction. The Sele idealization is a reasonable approximation of the effect of an adjacent row of cells; correction of strong localizations of vertical magnetic field can stabilize cells, as can correction of metal pad current distributions that have become disturbed through deposits on or damage to cathodes.

Increased computational capability has made it possible to relax the high degree of idealization. Whereas the earlier work was oriented toward development of dispersion relationships that characterized the behavior of the least stable mode of oscillation, more recent work has been built on numerical extraction of eigenvalues of rather large matrices stemming from more comprehensive descriptions of the cells. A strong physical incentive for this change is that some mechanisms can lead to mode mixing in the response of the interface; even though the problem has been linearized, superposition of independent modes, each satisfying the governing equations, is not a valid approach. Ziegler [72] and Sneyd and Wang [73] showed the effect of the steady velocity and the magnetic field, respectively, in creating mode mixing. Shin and Sneyd [74], in a truncation of the system of equations at two modes, remarked that the interface instability was strongest when the frequency of two natural gravity waves (the waves arising from merely gravity and inertia after a disturbance of the interface) are nearly equal. They concluded that, since these frequencies are dependent on the geometry of the cell, there is an opportunity for designing the cell to avoid contributing to instability. A similar conclusion was expressed by Bojarevics and Romerio [75] in an extensive analysis that led to a generalization of the Sele criterion. In this generalization,

various unstable modes are linked to distinct Fourier components of the vertical magnetic field distribution, allowing more extensive, yet still practically simplified, application of this approach.

In more computationally oriented work, Segatz and Droste [76], have examined the dependence of instabilities on design and operating parameters. Figure 23 is representative of their results; the growth rate of various modes of oscillation is plotted against cell current. Below a certain current (approximately 57 kA in the present case) interface waves have essentially zero growth rate but beyond this current, instability occurs. Interestingly, the form of the unstable oscillation changes as the current is increased. In a similar way, but simplified by the use of the shallow water model (described in the previous section), Droste and colleagues [77]) calculated results not very different from those from the more complicated 3D model.

In general, results from these models are in agreement with operational experience. For example, both Bojarevics and Romerio [75] and Segatz and Droste [76] indicate that their calculated growth rates were less for deeper metal pools, larger anode–metal distances and lower currents, all consistent with industrial operations. In addition, Segatz and Droste compared their calculations against measurements on an actual cell (where oscillation of the interface below an anode shows up in the anode voltage) and the good match can be seen in Fig. 24. However, these studies are still based on considerable simplification, including that the steady velocity is zero. Both Shin and Sneyd [74] and Ziegler [72], on examining instabilities that include the steady velocity or Kelvin-Helmholtz effect, report that the effect of the anode–metal distance is not monotonic; under some conditions, increasing this distance can increase the instability. In a more comprehensive computational study, Descloux

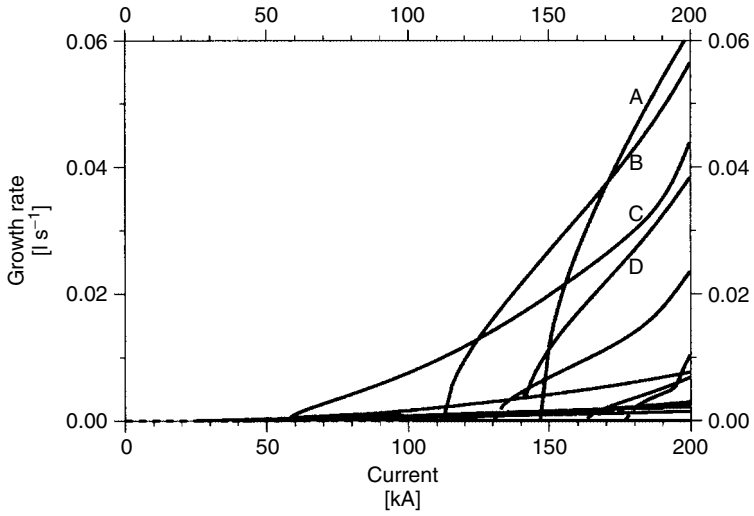


Fig. 23 Growth rates of interface instabilities computed by Segatz and Droste for various modes of oscillation of the metal–electrolyte interface as a function of current [76].

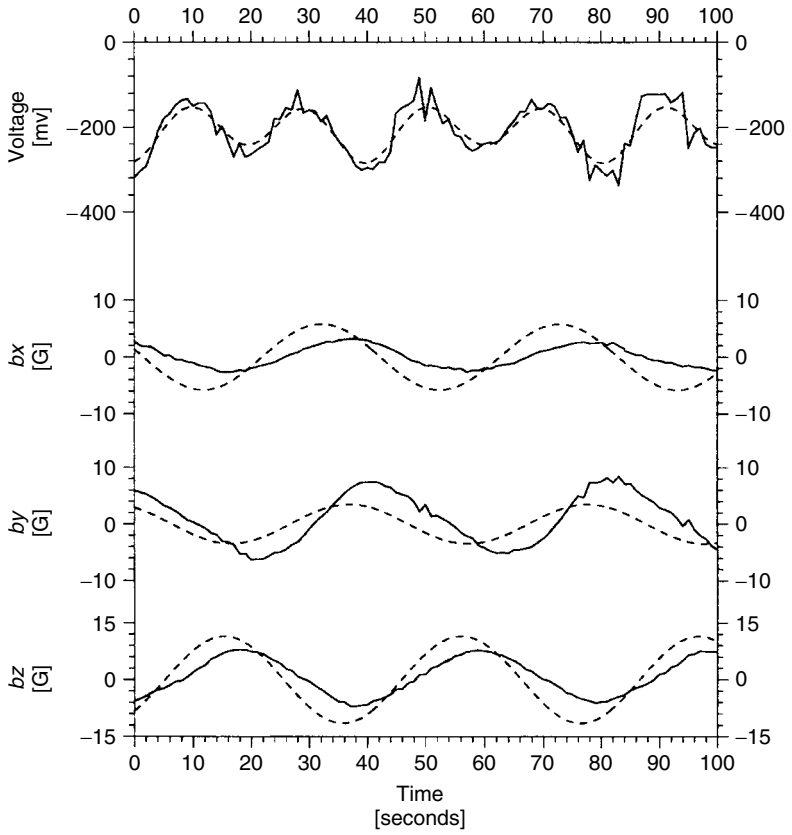


Fig. 24 The fluctuation of anode voltage and magnetic field due to metal–electrolyte interface oscillation as measured (solid line) and as computed by Segatz and Droste [76].

et al. [78] reported that the effect of the steady velocity was relatively weak in inducing instability.

Also in relation to the steady velocity, Antille et al. [79, 80] have shown how measurements of the anodic current fluctuations that arise from oscillation of the electrolyte–metal interface beneath the anode, can be used, in conjunction with their mathematical model, to deduce the steady velocity fields in a cell. The method has the advantage that there is no need to determine the electromagnetic forces, or to solve the Navier–Stokes equations and only a relatively small amount of computing is involved. They suggest that the method could also be used dynamically to follow the evolution of the velocity and have used it to compute waves in the electrolyte–metal interface.

In their studies of interface instability, most of the investigators listed above have employed a Fourier analysis of the coupling between the fluid motion and the electromagnetic body force. Davidson [81] and, in simpler form, Davidson and Lindsay [82] have pursued an alternative approach whereby a global energy balance

is used to assess stability. This analysis shows the role of the force transverse to the wave crest in providing energy to the disturbance, and provides an alternative physical illustration of the Sele mechanism. In a follow-on, the same authors (Davidson and Lindsay [83]) report that destabilization of modes with nearly identical gravity wave periods is not guaranteed, but rather depends on the coupling between the modes, which depends on the magnetic field distribution. Also, like Segatz and Droste, they report that at a certain current, the mode with the highest growth rate is not necessarily the one that first became unstable as the current was increased. Finally, Davidson [84] has proposed an active control system in which small anode movements are used to defeat incipient instabilities.

At the head of this section it was stated that an alternative to the fundamental modeling of idealized cells, just described, is the computationally intensive transient modeling of cells with a more faithful representation of cell features. This alternative has the advantage of dealing with more realistic situations, but may have the drawback of not successfully illuminating the root causes of disturbances.

One of the first papers, which is a representative of the latter approach is by Wahnsiedler [85]. This investigator used the commercial code ESTER® to simulate both the steady state and transient behavior of a 155 kA Alcoa cell. Cell details such as the ledge were included. The behavior of waves in the electrolyte–metal interface was computed and it was suggested that the wave motion was due to the interaction between the current and the varying magnetic field. Potocnik [86] and more recently Segatz et al. [87] also used ESTER in their similar studies of interface waves. While this approach is becoming increasingly feasible because of the increased availability of fast computation, the need to use small time steps is a continuing drawback.

Segatz [88] and coworkers have taken a hybrid approach to follow the evolution of the interface between the electrolyte and metal after a typical cell disturbance, the changing of an anode. The commercial software ESTER was used to compute the steady states before and after the anode change. Then the authors used the MHD analysis of Segatz and Droste (see above) to describe the trajectory of the cell between these two steady states. In this way, the anode-changing pattern for a cell could be optimized.

A fully transient simulation that does not use commercial software, but is based on the shallow water approximation relevant to Hall cells has been reported by Zikanov et al. [89]. Because of the speed of this simulation, additional features become available for investigation, such as continuing oscillations of constant amplitude, which can sometimes be detected in operating cells. This behavior, due to nonlinear interactions, is not within the capabilities of the linearized stability analyses that have formed the bulk of the literature to date.

4.2.5.3 Other Computational Models

An important new area in Hall cell simulation is that of dynamic heat and mass balance. These models simulate the interactions between changes in the cell heat balance and mass balance. For example, increased heat input can result in melting of some of the ledge material. This melting has two effects: it changes the rate of heat loss through the sidewall and also, depending on the ledge and bath compositions, changes the bath chemistry

and thus its melting point. In turn, these changes feed back on the bath temperature and superheat. All of these phenomena occur on somewhat different timescales, and enough of them are sufficiently slow so that, in practice, a new disturbance usually occurs to the cell before a new steady state is reached. Consequently, it is useful to model the behavior of the cell through various upsets and their denouements.

These models are useful in examining the effects of various perturbations such as anode changing, metal tapping, or changes in line current. They can be used to exercise on-line alumina controllers and algorithms and to help develop bath ratio control techniques. Finally, they can be useful training aids for plant engineers and pot operators.

An example is the model described by Tabsh and colleagues [90] with the results of calculations compared with measurements in an Alcan Brazil cell in Fig. 25. An additional case of good agreement of a model (including thermal phenomena) with experimental data is seen in Fig. 26 from the work of Tang and colleagues [91] at Kaiser Aluminum. These investigators compared the results of their model to measurements where a 70 kA prebaked cell was subjected to an addition of 40 kg of aluminum fluoride (changing the composition of the electrolyte and therefore unbalancing the equilibrium between electrolyte and ledge) or to an increase of 0.4 V in the applied voltage. The *long-term* dynamic behavior of the cell appears to be well simulated.

The models described above have all been transport phenomena models at the scale of the pot. Besides these, other models are finding application in the smelting area. At a somewhat finer scale, modeling has been used to examine the irregularity that develops on the bottom of an anode as it oxidizes. Even if the anode is homogeneous and flat, nonuniformity of the current density results in rounding of the initially sharp edges of the anodes. Zoric et al. [92] modeled the evolution of anode shape by performing a 2D computation of the current distribution in the electrolyte volume among an anode, ledge, and the aluminum. Then Faraday's law was used to follow the development of the anode shape with that shape reaching a

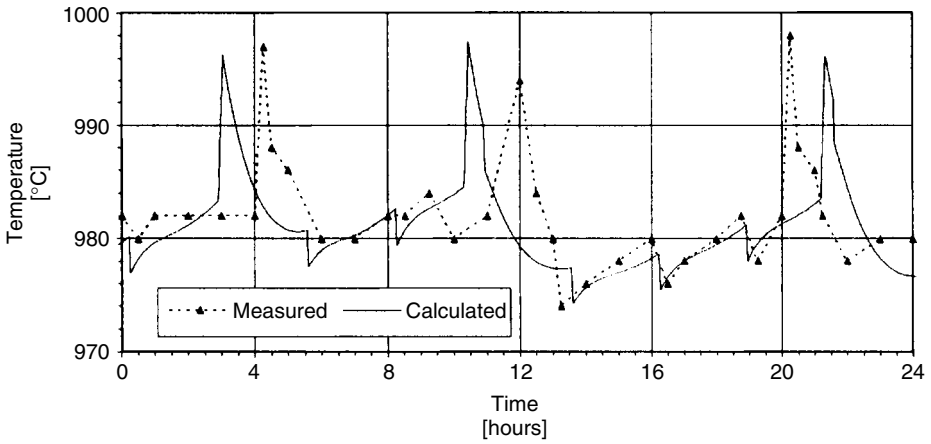


Fig. 25 Comparison of the results of the mathematical model of Tabsh and coworkers for the thermal behavior of a Hall–Héroult cell with measurements [90].

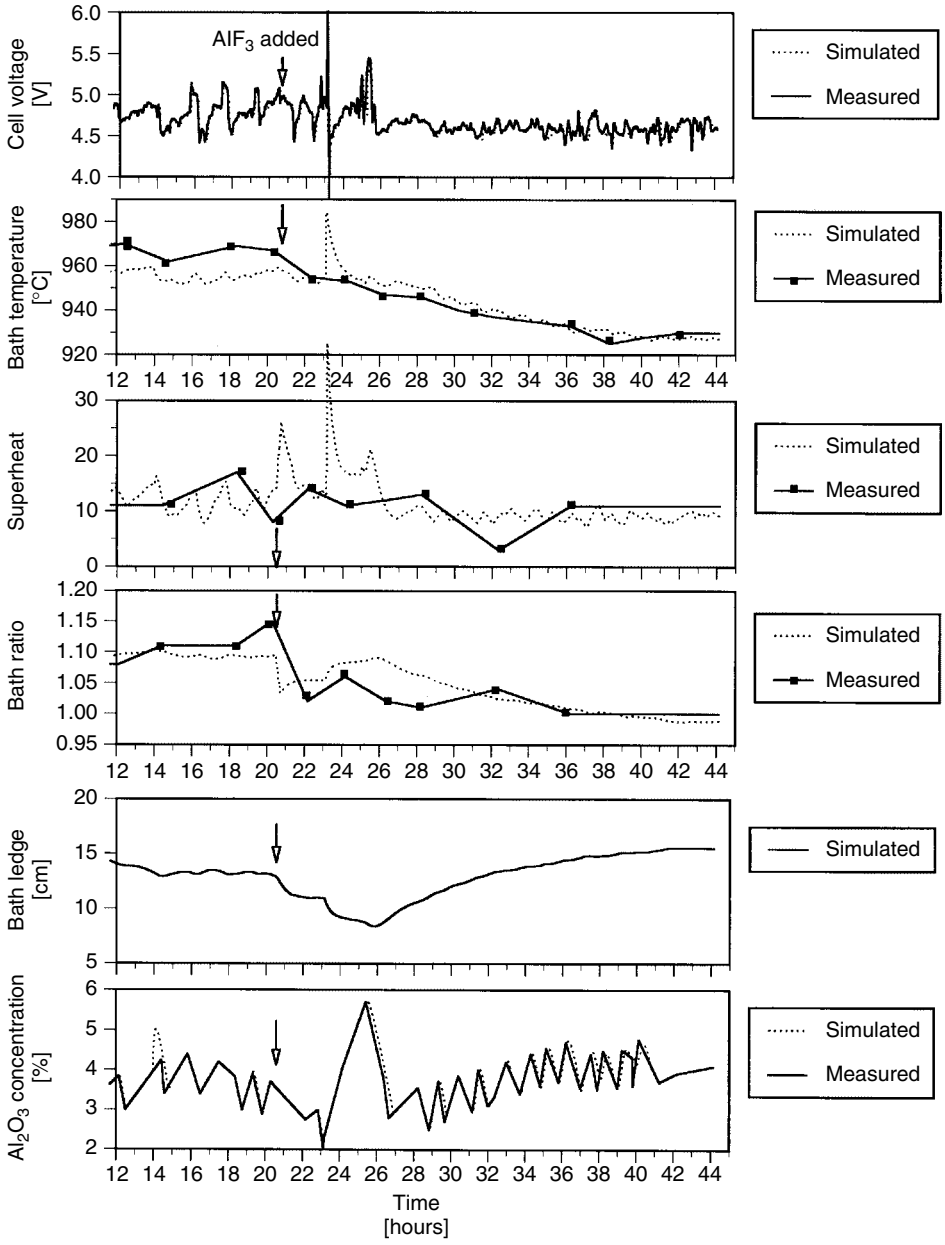


Fig. 26 Comparison of the predictions of the comprehensive mathematical model of Tang et al. with measurements for a Hall–Héroult cell, during a period when 40 kg of aluminum fluoride was added to the electrolyte [91].

steady state (assuming a steady downward movement of the whole anode) after a few days operation. Figure 27 is representative of their results for a cell where the ledge (to the left) was separated by 10 cm from the anode (to the right); the distance from the bottom of the anode to the aluminum (below) was 4.5 cm. Steady state was reached after about 8 days and the computed shape of the rounded edge of the anode matches very well with the measured shape of an anode removed from a cell.

Another area that has received some attention is the modeling of airflows exterior to the cells. This is relevant to the heat loss from the pots and to various environmental considerations; worker exposure to heat, chemicals, and emissions of fluoride and dust from buildings containing the potrooms. For example, Eick and Vogelsang [93] have recently modeled cooling fins fitted to the exterior of high amperage reduction cells. Potline ventilation is also the subject of a recent numerical study by Holt et al. [94]. Finally, the previously mentioned Bos and coworkers [52] have recently reviewed Aluminum Pechiney's mathematical modeling in the Hall–Héroult cell area, particularly regarding collector bars, potline ventilation and anode baking furnaces.

4.2.5.4 Physical Models

Physical modeling has proceeded apace with mathematical modeling. Indeed, in

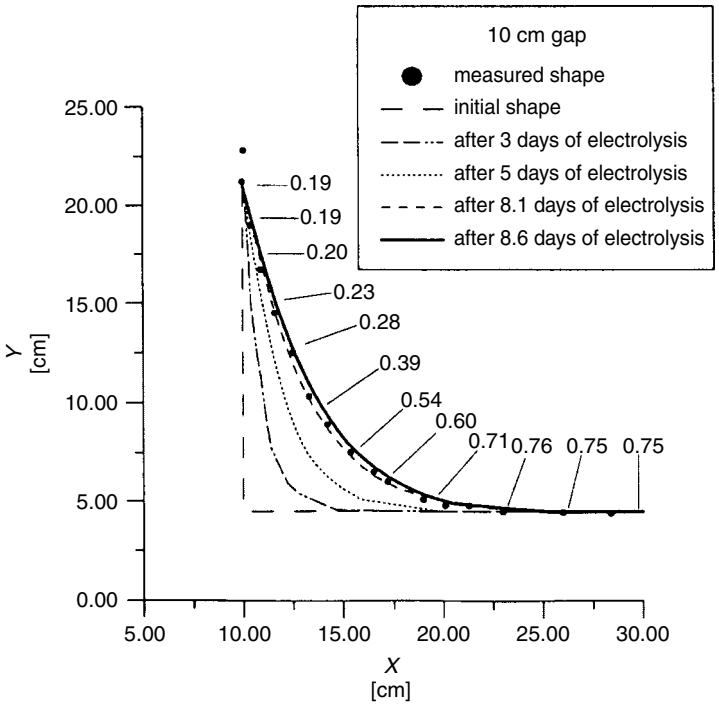


Fig. 27 Results of the computations of Zoric et al. for the evolution of anode shape over 8.6 days of electrolysis. The lower left of an anode is shown and the figure contains measurements of the final anode shape [92].

the 1970s Dernedde and Cambridge [95] used a physical model intended to simulate the flow driven by bubbles of gas from the anode. Electrolyte was simulated by water and aluminum by a dense organic phase consisting of a solution of perchlorethylene and silicone oil, with gas evolution being from porous, anode shaped, devices through which air was injected. The nature of the flow and the behavior of the electrolyte/metal interface were studied as a function of ledge position.

Physical models for bubble behavior and bubble driven flow have been described by Solheim and coworkers [96] who have also constructed mathematical models for this flow.

There has been much interest within the aluminum industry in "advanced" Hall–Héroult cells where the anode would be made of an inert material, rather than carbon. The anodic reaction is then merely the generation of oxygen, so the anode is not consumed. The geometry of the anode can then be more complicated than the simple blocks of the conventional anodes. A further advance is if the carbon "cathode" is replaced with one that is wetted by aluminum, such as a TiB_2 cathode. Both anode and cathode surfaces can then be inclined to facilitate gas removal. Efficient egress of gas bubbles is important because the bubbles, being nonconducting, increase the voltage difference between anode and cathode and therefore the energy consumption. Shekhar and Evans [44, 45] used water models to examine the flow in advanced cells, as well as, for comparison purposes, conventional cells. Velocities of the simulated electrolyte (water containing a small amount of butanol to achieve the correct bubble size) were measured in the gap between a simulated anode and cathode by laser-Doppler velocimetry. Figures 28 and 29 show the velocities (the space between electrodes is seen from above) for a flat horizontal anode (representing a conventional anode) and

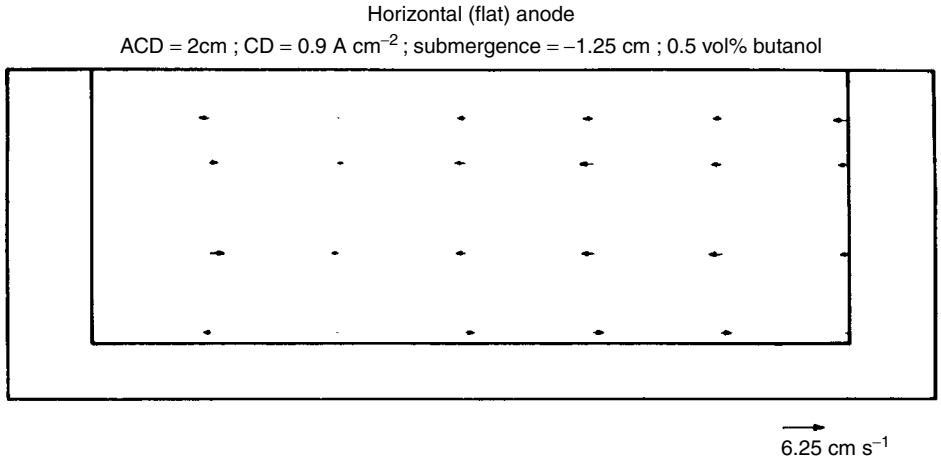


Fig. 28 Results of measurements, using laser-Doppler velocimetry, by Shekhar and Evans on a water model for studying gas driven electrolyte flow beneath anodes. Flow beneath a flat horizontal anode (inner rectangle), as seen from above, is depicted. The flow is seen to be very slow [44, 54].

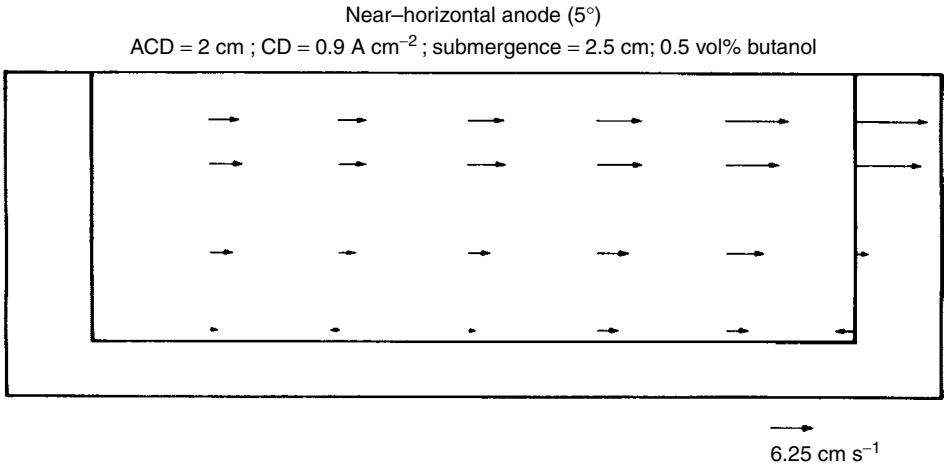


Fig. 29 Results similar to those of Fig. 28 but with the bottom of the anode inclined at 5° to the horizontal [44, 45].

for an anode with its lower surface inclined at 5° to the horizontal which is completely submerged (representing an anode in an advanced cell). Velocities are very much higher in the simulated advanced cell. These authors estimated the interelectrode resistance of advanced cells and therefore the likely energy saving.

Banerjee and Evans [97] constructed a one-twelfth scale physical model to study the electromagnetically driven flow in the Hall–Héroult cell. The cell structure and anodes were simulated with stainless steel and the only fluid in the cell was a low melting point alloy (Wood's alloy). Figure 30 depicts the physical model in cross section. With current densities comparable to those of an actual cell, the Lorentz forces were approximately an order of magnitude less. However, flows

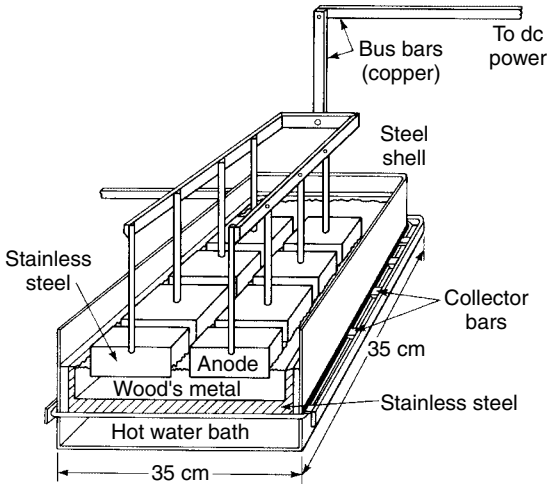


Fig. 30 The physical model of Banerjee and Evans for studying flow in Hall–Héroult cells driven by electromagnetic forces [97].

in the range of 1 cm s^{-1} were measured within the model using an electromagnetic velocity probe, as were the magnetic fields within and around the model (using a Hall effect probe). The effects of common cell disturbances, such as replacement of an anode, were examined with this model.

4.2.6 Novel Electrodes for Hall–Héroult Cells

The enormous electrical energy consumption of Hall–Héroult cells and their emissions of greenhouse gasses might be moderated by the use of very different anodes and cathodes. This has led to many investigations of inert anodes and "wettable" cathodes. The former electrodes are ones where the anodic reaction is the generation of oxygen and therefore the anode ought not to be consumed. This anodic reaction has equilibrium potential (versus an aluminum cathode) of 2.2–2.3 V compared to 1.7–1.8 V for a carbon anode evolving CO_2. Consequently, a switch to inert anodes would entail an *increase* in electrical energy consumption unless other changes are made. One such additional change would be the reduction in the ACD thereby reducing the wasteful IR loss in conducting electricity through this distance containing relatively poorly conducting electrolyte intermixed with bubbles. This reduction in ACD is more likely to be possible with inert anodes because the location of the bottom of the anode is known more precisely than in the case of the consumable carbon anode. Consequently, the anode can be brought closer to the cathode without danger of short circuits. This reduction of the ACD can be further enhanced if materials wetted by aluminum are used in the bottom of the cell. The normal lining of the bottom of the cells is carbon, which is not wetted by aluminum. However, if this is replaced with a wetted material such as titanium diboride there is an opportunity for cell designs where only a thin liquid film of aluminum lies beneath the anodes. This film might drain to a much deeper sump, away from the anodes, from which it is practical to remove the metal with the usual siphon. With aluminum as a thin film the dangers of deformation and oscillation of the metal–electrolyte interface are eliminated and it might be possible to reduce the ACD to very small distances, perhaps of the order of a centimeter, with great saving in electrical energy consumption.

In the early 1980s an extensive investigation of inert anodes was carried out by Alcoa [98–101]. Much of the investigation focused on cermet anodes consisting of a nickel ferrite–nickel oxide ceramic phase and a nickel metal phase although, toward the end of the study it became apparent that a similar cermet, but with a copper alloy as the metal phase would be more resistant to the hostile conditions of the Hall–Héroult cell. Cermet materials appear promising in this application because they offer higher electrical conductivity and thermal shock resistance than ceramics but are likely to be less prone to anodic oxidation than metals. The investigation at Alcoa included extensive measurements of the solubility and dissolution kinetics of various metal oxides in baths of various compositions and over a range of temperatures [98]. Nickel ferrite was found to be less soluble than either nickel oxide or Fe_2O_3. Linear sweep voltammograms of cermet anodes with nickel as the metal phase showed significant currents at potentials below the potential at which oxygen evolution should occur [99]. With a copper alloy as the metal phase such currents were much lower, perhaps nonexistent. Bench scale Hall–Héroult cells were run

using inert anodes and showed some contamination of the aluminum product by nickel and iron and higher corrosion rates/contamination if the alumina content of the bath was not kept close to saturation [100]. Inert anodes with a copper metal phase performed much better in these bench scale tests.

A recent investigation of the $NiO-NiFe_2O_4-Cu$ cermet material, and of tin oxide anode material, is that of Xiao and coworkers [102]. Figure 31 is taken from their work and shows the corrosion rate of the former material at cell current/temperature as a function of alumina content. The densities of these materials were 6.0–6.3 g cm^{-3}. Consequently a corrosion rate of 0.05 *g* cm^{-2} h^{-1} means that the material is disappearing at the rate of 0.08 mm h^{-1}. This means that the material is hardly inert, for example, a reasonable thickness of 10–20 cm would be consumed in a few months. However, this number should be contrasted with the oxidation rate of a carbon anode at this current density, which is of order 0.5 mm h^{-1}. There are reports [103, 104] that protective films of cerium oxide, formed in situ, can reduce the corrosion rate of inert anodes of this type. Sadoway [105] has suggested that metal alloys that form protective coatings could be practical inert anodes.

Kvande [106] has reviewed recent work on inert anodes (and wettable cathodes).

In the early 1980s Kaiser Aluminum investigated the use of titanium diboride as a wettable cathode material for Hall–Héroult cells [47]. Similar investigations by Reynolds Metals Co. continued until that company's recent merger with Alcoa [107]. The Reynolds work, and earlier research and development by Martin Marietta Aluminum [108], involved TiB_2-C composites. Approaches to the

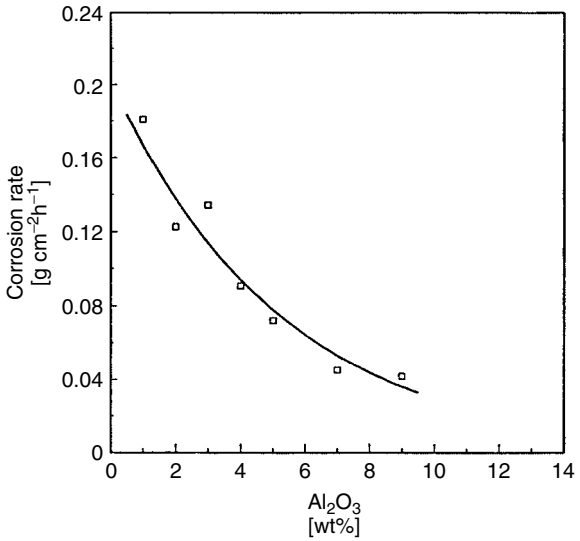


Fig. 31 Measurements of the rate of corrosion of a $NiO-NiFe_2O_4-Cu$ cermet anode material as a function of alumina concentration in the electrolyte. From the work of Xiao et al. [102].

use of these cathode materials fall into two rough categories:

1. placing (e.g. by tiling or troweling) a layer of wettable material onto the surface of a conventional carbon cathode
2. using solid objects (e.g. mushroom shapes) placed into the bottom of cells.

Both approaches are suitable for retrofitting existing cells (the former with the incorporation of a sump or well into some part of the cell where the metal may accumulate) and both complement the use of inert anodes. However, the latter approach introduces the possibility of radically different designs for Hall–Héroult cells such as the vertical electrodes examined by Dorward and Payne [47] or the bipolar design suggested by LaCamera [109]). Shekhar and Evans [44–46] have modeled some of these designs and provided estimates of potential energy savings resulting from their use.

4.2.7 Concluding Remarks

This article has described the Hall–Héroult cell that is the mainstay of the aluminum industry throughout the world. Emphasis has been on the electrochemistry and electrochemical engineering that govern cell performance. The cell operation, electrolyte chemistry, thermodynamics, and electrode kinetics have been reviewed. Some complexities, notably the ''anode effect'' and the environmentally important fluoride emissions and anode gas bubbles and their effect on cell voltage, flow, and CE, have been examined. The incorporation of these phenomena, along with current distribution, magnetic fields, electromagnetically driven flow, heat and mass transport, and cell instability into mathematical models was summarized. Mention was made of physical modeling of Hall–Héroult cells and of the possibility of radical improvement in cell performance resulting from new materials for inert anodes and wettable cathodes.

Acknowledgment

One of the authors (JWE) acknowledges the support received, over many years, for his work on aluminum from the US Department of Energy, Bureau of Mines, Alcoa, and Reynolds Metals Corp.

References

1. International Aluminium Institute at www.world-aluminium.org.
2. J. W. Evans, *Metall. Mater. Trans.* **1995**, *26*, 189–208.
3. P. T. Stroup, *Trans. TMS-AIME* **1964**, *230*, 356–371.
4. R. A. Frank, C. W. Finn, J. F. Elliott, *Metall. Mater. Trans. B* **1989**, *20*, 161–173.
5. B. Wu, R. G. Reddy, R. D. Rogers, in *Light Metals 2001* (Ed.: J. L. Angier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, pp. 237–243.
6. A. S. Russell, *Metall. Mater. Trans. B* **1981**, *12*, 203–215.
7. K. Grjotheim, C. Krohn, M. Malinovský et al., *Aluminium Electrolysis: The Chemistry of the Hall-Héroult Process*, Aluminium-Verlag, Düsseldorf, 1977.
8. K. Grjotheim, B. J. Welch, *Aluminum Smelter Technology*, Aluminium-Verlag, Düsseldorf, 1980.
9. K. Grjotheim, H. Kvande (Eds.), *Understanding the Hall-Héroult Process for Production of Aluminum*, Aluminium-Verlag, Düsseldorf, 1986.
10. A. R. Burkin (Ed.), *Production of Aluminium and Alumina*, Critical Reports on Applied Chemistry (Society of Chemical Industry), Vol. 20, John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1987.
11. W. S. Peterson (Ed.), *Hall-Héroult Centennial*, The Minerals, Metals and Materials Society, Warrendale, PA, 1986.

12. J. W. Evans, L. C. De Jonghe, *The Production and Processing of Inorganic Materials*, TMS, Warrendale, PA, USA, 2002.
13. F. McGravie, W. W. Robertson, S. T. Solinski, in *25th Conference on Metallurgists* (Eds.: E. Ozberk, D. W. Macmillan, R. I. L. Guthrie), The Metallurgical Society of the Canadian Institute of Mining and Metallurgy, Series 25-7, No. 2, 1986, pp. 3–19.
14. C. Vanvoren, P. Homsi, J. L. Basquin, in *Light Metals 2001* (Ed.: J. L. Angier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, pp. 221–226.
15. E. Skybakmoen, A. Solheim, A. Sterten, *Metall. Mater. Trans. B* **1997**, *28*, 81–86.
16. W. Haupin, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 531–537.
17. W. Haupin, H. Kvande, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 379–384.
18. J.-H. Yang, Q.-X. Zhang, J. Thonstad, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 453–460.
19. H. Zhu, D. R. Sadoway, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 241–246.
20. T. S. Sørenson, S. Kjelstrup Sadoway, in *Light Metals 1999* (Ed.: C. E. Eckert, The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 415–422.
21. L. N. Solli, Personal communication, May 26, 1998.
22. T. G. Pearson, J. Waddington, *Discuss. Faraday Soc.* **1947**, *1*, 307–320.
23. M. M. R. Dorreen, M. M. Hyland, B. J. Welch, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 483–489.
24. A. Kisza, J. Kazmierczak, J. Thonstad, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 423–429.
25. J. Marks, R. Roberts, V. Bakshi, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 365–371.
26. A. T. Tabereaux, *J. Metals* **1994**, *11*, 32–34.
27. J. Thonstad, T. A. Utigard, H. Vogt Dolin, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 249–236.
28. J. Marks, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 287–291.
29. J. Marks, A. Taberaux, D. Pape et al., in *Light Metals 2001* (Ed.: J. L. Angier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, pp. 295–302.
30. E. J. Dolin, J. Casola, T. Miller, in *Light Metals 2001* (Ed.: J. L. Angier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, pp. 319–325.
31. H. Zhu, D. R. Sadoway, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 257–263.
32. H. Vogt, *Metall. Mater. Trans. B* **2000**, *31*, 1225–1230.
33. M. M. R. Dorreen, D. L. Chin, J. K. C. Lee, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 311–316.
34. J. L. Henry, R. D. Holliday, *J. Metals* **1957**, *9*, 1384–1385.
35. R. J. Moreau, D. Ziegler, *Light Metals 1986*, The Minerals, Metals and Materials Society, Warrendale, PA, 1986, pp. 199–1211.
36. I. Utne, K. A. Paulsen, J. Thonstad, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 293–301.
37. M. M. Hyland, B. J. Welch, J. B. Metson, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 333–338.
38. F. Hine, *Electrode Processes and Electrochemical Engineering*, Plenum Press, New York, 1985.
39. S. Fortin, M. Gerhardt, A. J. Gesing, in *Light Metals 1984* (Ed.: J. P. McGeer), The Minerals, Metals and Materials Society, Warrendale, PA, 1994, pp. 721–741.
40. T. M. Hyde, B. J. Welch, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, pp. 333–340.
41. R. J. Aaberg, V. Ranum, K. Williamson, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, pp. 341–346.
42. X. Wang, A. T. Tabereaux, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals,

Metals and Materials Society, Warrendale, PA, 2000, pp. 239–247.
43. J. Xue, H. A. Øye, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 247–253.
44. R. Shekhar, J. W. Evans, *Metall. Mater. Trans. B* **1994**, *25*, 333–340.
45. R. Shekhar, J. W. Evans, *Metall. Mater. Trans. B* **1994**, *25*, 341–347.
46. R. Shekhar, J. W. Evans, *Metall. Mater. Trans. B* **1996**, *27*, 19–27.
47. R. C. Dorward, J. R. Payne, Report DOE/CS/40215-2, Kaiser Aluminum and Chemical Corporation, June 1985.
48. E. D. Tarapore, *Light Metals 1979*, The Metallurgical Society of AIME, New York, NY, 1979, pp. 541–550, Vol. 1.
49. J. W. Evans, Y. Zundelevich, D. Sharma, *Metall. Mater. Trans. B* **1981**, *12*, 353–360.
50. S. D. Lympany, J. W. Evans, *Metall. Mater. Trans. B* **1983**, *14*, 63–70.
51. V. Potocnik, J. W. Evans, in *25th Conference of Metallurgists* (Eds.: E. Ozberk, D. W. Macmillan, R. I. L. Guthrie), The Metallurgical Society of the Canadian Institute of Mining and Metallurgy, Series 25-7, No. 2, Montreal, Québec, 1986, pp. 23–38.
52. J. Bos, G. Bouzat, J. Colin de Verdiere et al., in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 393–401.
53. M. Dupuis, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 297–302.
54. S. A. Sherbinin, V. V. Pingin, A. G. Barantsev, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 323–329.
55. D. P. Ziegler, R. L. Kozarek, in *Light Metals 1991* (Ed.: E. Roy), The Minerals, Metals and Materials Society, Warrendale, PA, 1991, pp. 381.
56. L. Guohua, L. Dexiang, in *Light Metals 1995* (Ed.: J. W. Evans), The Minerals, Metals and Materials Society, Warrendale, PA, 1995, pp. 301–303.
57. M. Segatz, D. Vogelsang, in *Light Metals 1991* (Ed.: E. Roy), The Minerals, Metals and Materials Society, Warrendale, PA, 1991, pp. 393.
58. R. F. Boivin, *A Volume-Integral Method for Calculating the Magnetic Field in Aluminum Electrolysis Cells*, Report No. R95-50, CERCA, Montreal, April 11, 1995.
59. A. Kalimov , V. Krukovski, L. Minevich et al., *Abstract of 14th International Conference on Magnetohydrodynamics*, Riga, Latvia, 1995, p. 124.
60. S. Yang et al., in *Light Metals 2001* (Ed.: J. Anjier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, p. 433.
61. V. Potocnik, F. Laroche, in *Light Metals 2001* (Ed.: J. Anjier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, p. 419.
62. K. J. Fraser, M. P. Taylor, A. M. Jenkin, in *Light Metals 1991* (Ed.: C. M. Bickert), The Minerals, Metals and Materials Society, Warrendale, PA, 1990, p. 221.
63. K. Bech, S. T. Joansen, A. Solheim, in *Light Metals 2001* (Ed.: J. Anjier), The Minerals, Metals and Materials Society, Warrendale, PA, 2001, p. 463.
64. K. Kalgraf, K. Tørklep, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 455–464.
65. K. Tørklep, K. Kalgraf, T. Nordbø, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, p. 377–386.
66. R. J. Moreau, J. W. Evans, *J. Electrochem. Soc.* **1984**, *131*, 2251–2259.
67. R. J. Moreau, D. Ziegler, *Metall. Mater. Trans. B* **1988**, *19*, 737–744.
68. T. Sele, *Metall. Mater. Trans. B* **1977**, *8*, 613–618.
69. N. Urata, *Light Metals 1985*, The Minerals, Metals and Materials Society, Warrendale, PA, 1985, pp. 581–589.
70. D. Sneyd, *J. Fluid Mech.* **1985**, *156*, 223–236.
71. D. Sneyd, *J. Fluid Mech.* **1992**, *236*, 111.
72. D. P. Ziegler, *Metall. Mater. Trans. B* **1993**, *24*, 899.
73. A. D. Sneyd, A. Wang, *J. Fluid Mech.* **1994**, *263*, 343.
74. D. Shin, A. D. Sneyd, in *Light Metals 2000* (Ed.: R. D. Peterson), The Minerals, Metals and Materials Society, Warrendale, PA, 2000, pp. 279–283.
75. V. Bojarevics, M. V. Romerio, *Eur. J. Mech., B/Fluids* **1994**, *13*, 33–56.

76. M. Segatz, C. Droste, in *Light Metals 1994* (Ed.: U. Mannweiler), The Minerals, Metals and Materials Society, Warrendale, PA, 1994, pp. 313–322.
77. Ch. Droste, M. Segatz, D. Vogelsang, in *Light Metals 1998* (Ed.: B. Welch, The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 419–428.
78. J. Descloux, M. Flueck, M. V. Romerio, in *Light Metals 1994* (Ed.: U. Mannweiler), The Minerals, Metals and Materials Society, Warrendale, PA, 1994, pp. 275–281.
79. J. P. Antille, M. Flueck, M. V. Romerio, in *Light Metals 1994* (Ed.: U. Mannweiler), The Minerals, Metals and Materials Society, Warrendale, PA, 1994, pp. 305–312.
80. J. P. Antille, J. Descloux, M. Flueck, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 333–338.
81. P. A. Davidson, *Eur. J. Mech., B/Fluids* **1994**, *13*, 15–32.
82. P. A. Davidson, R. I. Lindsay, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, pp. 437–442.
83. P. A. Davidson, R. I. Lindsay, *J. Fluid Mech.* **1998**, *362*, 273.
84. P. A. Davidson, R. I. Lindsay, *Mater. Sci. Technol.* **2000**, *16*, 475.
85. W. E. Wahnsiedler, in *Light Metals 1987* (Ed.: R. D. Zabreznik), The Minerals, Metals and Materials Society, Warrendale, PA, 1987, pp. 269–287.
86. V. Potocnik, in *Light Metals 1989* (Ed.: P. G. Campbell), The Minerals, Metals and Materials Society, Warrendale, PA, 1989, pp. 227–235.
87. M. Segatz, D. Vogelsang, C. Droste, in *Light Metals 1993* (Ed.: S. K. Das), The Minerals, Metals and Materials Society, Warrendale, PA, 1993, pp. 361.
88. M. Segatz, Ch. Droste, D. Vogelsang, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, pp. 429–435.
89. O. Zikanov, A. Thess, P. A. Davidson, et al.,*Metall. Mater. Trans. B* **2000**, *31*, 1541.
90. Tabsh, M. Dupuis, A. Gomes, *Light Metals 1996* (Ed.: W. Hale), The Minerals, Metals and Materials Society, Warrendale, PA, 1996, pp. 451–457.
91. H. Q. Tang, N. Urata, C. M. Read, in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 349–357.
92. J. Zoric, I. Rousar, J. Thonstad, in *Light Metals 1997* (Ed.: R. Huglen), The Minerals, Metals and Materials Society, Warrendale, PA, 1997, pp. 449–456.
93. I. Eick, D. Vogelsang, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, 1999, Warrendale, PA, pp. 339–345.
94. N. J. Holt, N. M. Anderson, M. Karlsen, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 263–267.
95. E. Dernedde, E. L. Cambridge, Gas induced circulation in an aluminum reduction cell, Paper presented at the *AIME Annual Meeting*, New York, 1975.
96. A. Solheim, S. T. Johansen, S. Rolseth, in *Light Metals 1989* (Ed.: P. G. Campbell), The Minerals, Metals and Materials Society, Warrendale, PA, 1988, pp. 245–252.
97. S. K. Banerjee, J. W. Evans, *Metall. Mater. Trans. B* **1990**, *21*, 59–69.
98. D. H. DeYoung, in *Light Metals 1986* (Ed.: R. E. Miller), The Minerals, Metals and Materials Society, Warrendale, PA, 1986, pp. 299–307.
99. G. P. Tarcy, in *Light Metals 1986* (Ed.: R. E. Miller), The Minerals, Metals and Materials Society, Warrendale, PA, 1986, pp. 309–320.
100. S. P. Ray, in *Light Metals 1986* (Ed.: R. E. Miller), The Minerals, Metals and Materials Society, Warrendale, PA, 1986, pp. 287–298.
101. J. D. Weyand, in *Light Metals 1986* (Ed.: R. E. Miller), The Minerals, Metals and Materials Society, Warrendale, PA, 1986, pp. 321–339.
102. H. Xiao, R. Hovland, S. Rolseth et al., *Metall. Mater. Trans. B* **1996**, *27*, 185–193.
103. C. F. Windisch,Jr., D. M. Strachan, C. H. Henager Jr et al., in *Light Metals 1993* (Ed.: S. K. Das), The Minerals, Metals and Materials Society, Warrendale, PA, 1993, pp. 445–454. Ed.:
104. J. S. Gregg, M. S. Frederick, H. L. King, in *Light Metals 1993* (Ed.: S. K. Das), The Minerals, Metals and Materials Society, Warrendale, PA, 1993, pp. 455–464.
105. D. R. Sadoway, *Light Metals 1990*, The Minerals, Metals and Materials Society, Warrendale, PA, 1990, pp. 403–407.

106. H. Kvande, in *Light Metals 1999* (Ed.: C. E. Eckert), The Minerals, Metals and Materials Society, Warrendale, PA, 1999, pp. 369–376.
107. A. Taberaux, J. Brown, I. Eldridge et al., in *Light Metals 1998* (Ed.: B. Welch), The Minerals, Metals and Materials Society, Warrendale, PA, 1998, pp. 257–264.
108. A. V. Cooke, W. M. Bucht, in *Light Metals 1985* (Ed.: H. O. Bohner), The Minerals, Metals and Materials Society, Warrendale, PA, 1985, pp. 545–566.
109. A. F. LaCamera, U.S. Patent 5,279,715, ALCOA, Pittsburgh, PA, 1994.

5
Electrowinning of Other Inorganic Compounds

Carl Heinz Hamann
Institut für Reine und Angewandte Chemie, Oldenburg, Germany

Thorsten Röpke
Hydrogen Institute of Applied Technologies, Schwerin, Germany

Peter Schmittinger
Formerly Degussa AG, Germany

5.1 Introduction 269

5.2 Chlorine and Alkali from Sodium Chloride, Chlorine from Hydrochloric Acid, Chlorine–Oxygen Compounds [1–7] 269
5.2.1 Amounts and Uses of Products 269
5.2.2 Basic Electrochemical Reactions 270
5.2.3 Chlorine and Alkali from Sodium Chloride 272
5.2.3.1 Amalgam Process 272
5.2.3.2 Diaphragm Process 273
5.2.3.3 Membrane Process 276
5.2.3.4 Treatment of Educts and Products 278
5.2.3.4.1 Brine 278
5.2.3.4.2 Chlorine 280
5.2.3.4.3 Caustic Alkali 280
5.2.3.5 Comparison of the Three Processes 281
5.2.3.6 Electrode Kinetics 283
5.2.3.6.1 Amalgam Cell 284
5.2.3.6.2 Amalgam Decomposition 286
5.2.3.6.3 Diaphragm and Membrane Cell 287
5.2.3.7 Closing Remarks 287
5.2.3.7.1 Finite Gap and Zero Gap Principle 287
5.2.3.7.2 Oxygen Depolarized Cathode (ODC) [25] 288
5.2.3.7.3 Production of Alcoholates, Polyolates, and Other Chemicals from Alkali Amalgam 290
5.2.4 Chlorine from Hydrochloric Acid 291

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

5.2.5	Chlorine–Oxygen Compounds [1, 4, 9, 10]	292
5.2.5.1	Hypochlorite	293
5.2.5.2	Chlorite	293
5.2.5.3	Chlorate	294
5.2.5.4	Perchlorates	295
5.2.5.5	Chlorine Dioxide	296
5.2.6	Additional Aspects	296
5.3	**Fluorine [42, 43]**	**297**
5.3.1	Introduction	297
5.3.2	Electrochemical Reactions	297
5.3.3	Technical Process	297
5.4	**Hydrogen, Oxygen, and Ozone [7, 48–51]**	**299**
5.4.1	Introduction	299
5.4.2	Electrochemical Reactions	300
5.4.3	Technical Processes	301
5.4.3.1	Classical Water Electrolysis	301
5.4.3.2	Membrane Electrolyzers	302
5.4.3.3	Steam Electrolyzers	304
5.4.3.4	Closing Remarks	305
5.5	**Peroxodisulfates [7, 56, 57]**	**305**
5.5.1	Introduction	305
5.5.2	Electrochemical Reactions	305
5.5.3	Technical Processes	306
5.6	**Hydrogen Peroxide [7, 57, 65]**	**307**
5.6.1	Introduction	307
5.6.2	Electrochemical Reactions	307
5.6.3	Technical Process	307
5.7	**Further Inorganic Species [71–73]**	**308**
5.7.1	Chromic Acid Solutions	309
5.7.2	Manganese Dioxide	309
5.7.3	Permanganate	310
5.7.4	Copper Oxide	312
	References	312

5.1 Introduction

In this chapter, roughly one dozen quite different production processes will be summarized. For instance, the most important electrochemical process, the electrolysis of aqueous sodium chloride leading to chlorine, alkali, hydrogen, and chlorine oxygen compounds are incorporated (Sect. 5.2). Other processes are carried out in a small scale only, but have an unchallenged place in the technical chemistry, for instance, the generation of molecular fluorine (Sect. 5.3). Others are of limited importance nowadays, but may gain importance in the future, for example, the electrolysis of water (Sect. 5.4).

As will be seen, in most cases electrochemically produced inorganic species are the outcome of anodic oxidations. This is due to the fact that this process offers a very high oxidation power. So it is used in technical processes if chemical oxidants with enough power do not exist (as in the case of the preparation of molecular fluorine) or if the chemical oxidants are too expensive or too difficult to prepare.

The texts are arranged in the following manner: amounts and uses of products – electrochemical reactions – technical process(es). In the case of Sect. 5.2, there are additional subsections concerning, for instance, the treatment of educts and products, the comparison of different possible processes or electrode kinetics. In Sects 5.3–5.6, such information, if given, are incorporated into the main points. Section 5.7 gives only rough information on the subjects incorporated. A more detailed description would go beyond the scope of this encyclopaedia. So, a selection had to be made. This holds for the number of references, too. Please note that inorganic electrochemistry in general is dealt with in Volume 7 of this encyclopaedia.

5.2 Chlorine and Alkali from Sodium Chloride, Chlorine from Hydrochloric Acid, Chlorine–Oxygen Compounds [1–7]

5.2.1 Amounts and Uses of Products

In the year 2004, the global demand of chlorine was 48.4×10^6 metric tons, more than 95% thereof by the electrochemical means described as follows. The percentage of world chlorine consumption in 2004

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

was as follows [8a]:

Vinyl chloride	34%
Organics	20%
Chlorinated hydrocarbons	6%
Water treatment	6%
Pulp and paper bleaching	4%
Other (inorganic products etc.)	30%

Considering the alkali, the worldwide demand of NaOH, caustic soda, was about 46×10^6 tons per year in 2002, again more than 90% by electrochemical means. The use in 2002 was as follows [8b]:

Production of organic chemicals	18%
Pulp and paper	18%
Inorganic chemicals	15%
Soap and detergents, textiles	12%
Alumina	8%
Water treatment	5%
Other	24%

The worldwide production of KOH, caustic potash, was $700–800 \times 10^3$ tons per year, almost exclusively by electrochemical means. Pure quality potassium hydroxide is used as a raw material for the chemical and pharmaceutical industry, in the dye synthesis, for photography as a developer alkali, as an electrolyte in galvanic elements/fuel cells and in the electrolysis of water (see Sect. 5.4). Technical quality KOH is used as a raw material in the detergent and soap industry, as a starting material for organic and inorganic potassium compounds, for the manufacturing of cosmetics, glass and textiles, and last but not least as a drying agent and CO_2-absorbent [9].

Chlorine–oxygen compounds discussed here are hypochlorite ClO^-, chlorite ClO_2^-, chlorate ClO_3^-, perchlorate ClO_4^-, and chlorine dioxide ClO_2. With the exception of chlorite, all chlorine–oxygen compounds mentioned can be obtained as a direct product of electrolysis, but electrolytic production is not dominating in all cases. There are chemical pathways, too, using molecular chlorine and caustic alkali as basic educts – which have been produced electrolytically before.

Chlorine–oxygen compounds are used – besides chlorine itself – predominantly for bleaching and disinfection purposes. With water they form solutions of hypochloric acid, these solutions showing oxidizing, chlorinating, or hydrochlorinating properties, depending on the pH-value established [10].

Since it became clear that the direct use of chlorine in the bleaching of pulp and paper leads to highly poisonous chlorinated organic species in the waste water, chlorine dioxide is used instead in the so-called ECF (Elemental Chlorine Free) process. The competing process is the TCF (Total Chlorine Free) process, which uses peroxides [11].

Chlorine dioxide is produced from sodium chlorate in situ, the latter becoming the most important technical chlorine–oxygen compound in this way (production in the year 2002: 2.8×10^6 tons). Ninety-five percent of the chlorate is used in the described way; another usage is, for instance, the production of herbicides.

Perchlorate as a strong oxidant is widely used. The usage in pyrotechnics is only one example.

5.2.2 Basic Electrochemical Reactions

If two electrodes are immersed into a sodium chloride brine and an electrolysis

is performed, a lot of electrode reactions are possible from the thermodynamical view. However, considering the activation overpotentials for different possible processes at different electrode materials, we obtain the following electrode reactions (for more information, see Sect. 5.2.3.6; this holds for the side reactions mentioned in the following text).

Using carbon or, better, a titanium surface activated with a mixture of titanium and ruthenium oxides as an anode material (the latter is the so-called *dimensionally stable anode* (DSA) [7, 12]), chlorine ions are discharged

$$2NaCl \longrightarrow Cl_2 + 2Na^+ + 2e^- \quad (1)$$

Using mercury as cathode material, we get sodium amalgam from Na^+

$$2NaCl + 2xHg + 2e^- \longrightarrow 2NaHg_x + 2Cl^- \quad (2)$$

Using iron as cathode material, hydrogen is split off from the water molecule

$$2H_2O + 2e^- \longrightarrow H_2 + 2OH^- \quad (3)$$

Considering reactions (1) and (2), chlorine and sodium amalgam are formed in the electrolysis cell. The amalgam can be extracted and undergoes a decomposition reaction in a separate reactor (decomposer, denuder). The electrocatalytic process is

$$2NaHg_x + 2H_2O \longrightarrow 2Na^+ + 2OH^- + 2xHg + H_2 \quad (4)$$

So, the addition of Eqs. (1), (2), and (4) gives the overall reaction of the so-called amalgam process for the production of chlorine and alkali from sodium chloride

$$2NaCl + 2H_2O \longrightarrow 2NaOH + H_2 + Cl_2 \quad (5)$$

As the products chlorine on the one hand and hydrogen and alkali on the other originate at different places, electrolysis cell and decomposer, there is no problem of separation.

Combining reaction (1) with (3) Eq. (5) is reproduced. But there are subsequent processes. First, the anodically generated chlorine will react quickly with hydroxide ions migrating from cathode to anode

$$3Cl_2 + 6OH^- \longrightarrow 3ClO^- + 3Cl^- + 3H_2O \quad (6)$$

The resulting hypochlorite then undergoes an anodic oxidation, forming chlorate

$$3ClO^- + \frac{3}{2}H_2O \longrightarrow ClO_3^- + 2Cl^- + 3H^+ + \frac{3}{4}O_2 + 3e^- \quad (7)$$

From chlorate, at the end, perchlorate can be produced anodically

$$ClO_3^- + H_2O \longrightarrow ClO_4^- + 2H^+ + 2e^- \quad (8)$$

(For reason of simplicity, we wrote H^+ instead of H_3O^+.)

So, the sodium hydroxide generated will be contaminated with chlorine–oxygen compounds, under rapid reduction of the current efficiency. The elemental chlorine bubbling from the anode will contain amounts of oxygen. Thus it is necessary to prevent OH^--migration from the cathode to the anode in order to receive an economic technical process. This can be done by using divided cells, leading to the so-called diaphragm and membrane processes.

On the other hand, if reactions (1) and (3) are performed in undivided cells, chlorine–oxygen compounds are accessible (see Sect. 5.2.5).

5.2.3 Chlorine and Alkali from Sodium Chloride

5.2.3.1 Amalgam Process

Mercury flows down the inclined base of the electrolytic cell (Fig. 1A). The base of the cell is electrically connected to the negative pole of the DC-supply. On the top of the mercury and flowing co-currently with it is a concentrated brine with a sodium content of ca. 310 g L^{-1} at the inlet. The brine must be purified thoroughly (see Sect. 5.2.3.4). Anodes are placed in the brine so there is a small gap between the anodes and the flowing mercury cathode. The anodes used nowadays are predominately of the DSA-type. They are constructed in form of parallel blades or rods, arranged in flow direction. The distance between these elements is needed for quick gas release.

The concentration of the amalgam is maintained at 0.2–0.4 wt% sodium (see Sect. 5.2.3.6). The chlorine gas and the depleted brine (270 g L^{-1}) flow out of the cell. The amalgam leaves the cell through a weir and flows into the decomposer.

There are horizontal decomposers (Fig. 1B) alongside or beneath the cell, or more often, nowadays, vertical decomposers or denuders (Fig. 1C). As mentioned in Sect. 5.2.2 (and described in detail in Sect. 5.2.3.6) the decomposition reaction is electrocatalytic, the catalyst being activated graphite. In the case of horizontal variety the catalyst is arranged in form of blades, 4–6-mm thick, which are partly immersed in the amalgam flowing down under influence of gravity. Demineralized water is pumped uphill above the amalgam surface and then is in contact with the blades, too. So, reaction (4) can take place, forming pure caustic alkali solution of 50 wt%.

In case of vertical variety, there is a packing of activated graphite spheres or lumps of 8–25 mm in diameter. The amalgam is fed in via an overhead distributor and trickles down in droplets through the bed of graphite and the

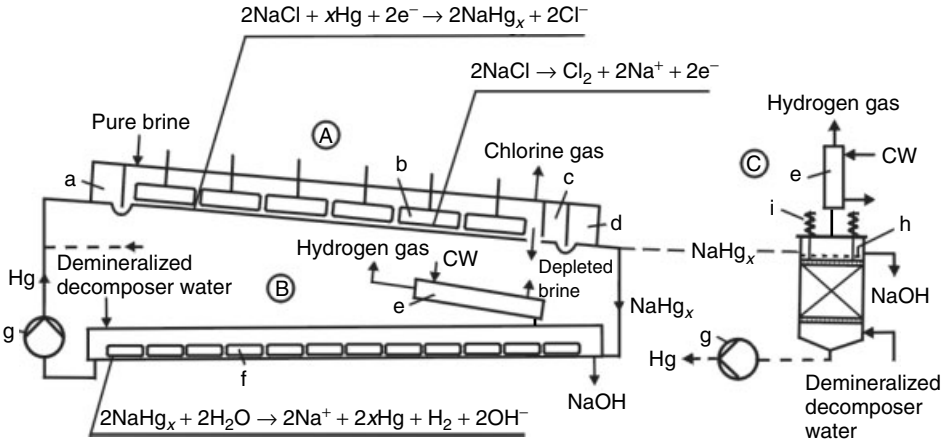


Fig. 1 Schematic view of a mercury cell with decomposers. (A) Mercury cell: (a) mercury inlet box; (b) anodes; (c) end box; (d) wash box. (B) Horizontal decomposer: (e) hydrogen gas cooler; (f) graphite blades; (g) mercury pump. (C) Vertical decomposer: (e) hydrogen gas cooler; (g) mercury pump; (h) mercury distributor; (i) packing pressing springs [3, p. 31].

water is pumped upwards counter to the amalgam. So, the electrocatalytic reaction (4) can take place again.

In both cases, the denuded mercury is pumped into the electrolysis cell again, as well as the brine, undergoing resaturation before.

A single amalgam cell is 0.8–2-m broad and has a length of 7–14 m. The total surface of the electrodes amounts to 30 m^2, the current density can be as high as 15 kA m^{-2}. This means 450 kA in total, corresponding to a production of 13 tons of chlorine per cell in a day.

The cell voltage under these conditions is ca. 4 V, the current efficiency 0.95. Up to 50 cells are connected in series. Horizontal decomposers usually have the same length as the cell, but are less broad. Vertical decomposers are packed 0.6–0.8-m high and have cross sections of 0.35 m^2 per 100 kA of cell current. The operating temperature is 75–90 °C for the electrolysis and 90–110 °C for the decomposition. Figure 2 shows a mercury cell with a vertical decomposer as an example.

Figures 3 and 4 show a mercury cell room from birds' eye view and a corresponding cross section. The cell room is roughly $70 \times 50 \times 10$ m in length, breadth, and height, respectively. The air must be changed 10–25 times per hour to remove the heat produced.

5.2.3.2 Diaphragm Process

The diaphragm cell uses a synthetic plastic or asbestos diaphragm surrounding the

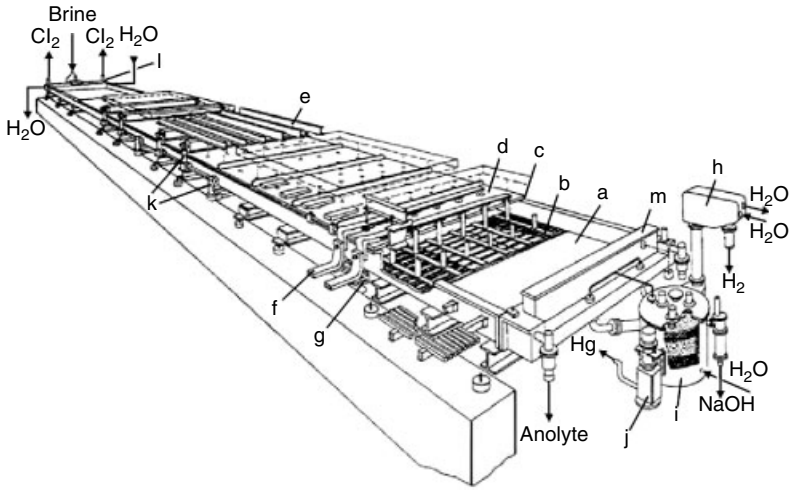


Fig. 2 Uhde mercury cell. (a) Cell base; (b) anode; (c) cover seal; (d) cell cover; (e) group adjusting gear; (f) intercell busbar; (g) short-circuit switch; (h) hydrogen cooler; (i) vertical decomposer; (j) mercury pump; (k) anode adjusting gear; (l) inlet box; (m) end box, compare Fig. 1. Please note: adjusting gears (e) and (k) are needed to maintain a distance of ca. 1.5–2.5 mm between anodes and the cathode. Greater distances lead to an increase of ohmic losses. In case of distances too small, the anode–cathode system will be short circuited and the cell will be damaged. In this case, the cell must be repaired and has to be short circuited in total by a suitable switch (g). In case of the anodes (b), the structure described in the text can be seen left, in the middle the substructure is displayed [3, p. 38, 13].

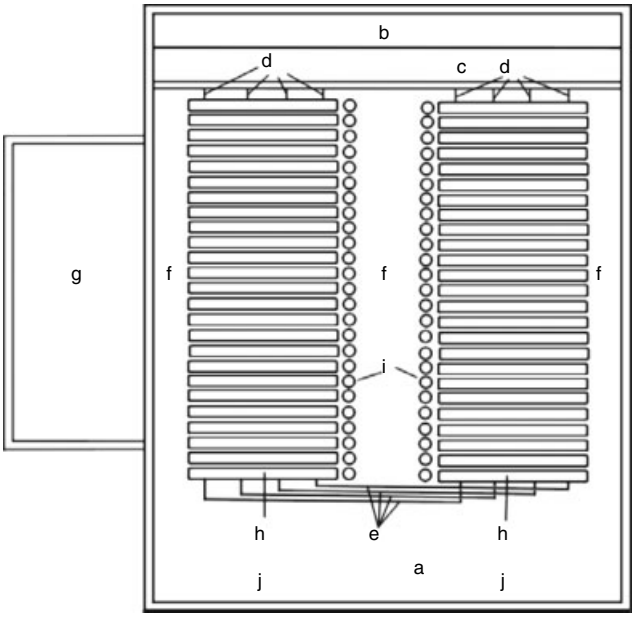


Fig. 3 Mercury cell room (bird's-eye view, schematic). (a) Cell room; (b) transformer room; (c) rectifier room; (d) busbars; (e) turn around busbars; (f) service walkways; (g) ancillary equipment; (h) electrolysis cells; (i) vertical decomposers; (j) cell assembly and maintenance area [3, p. 43].

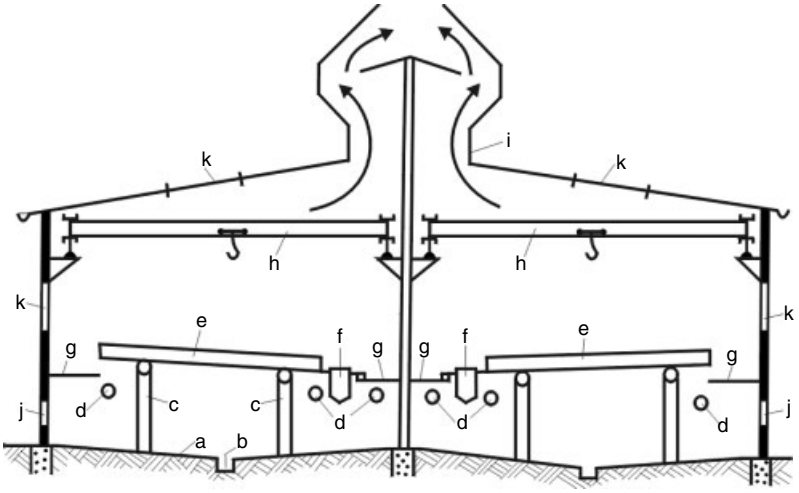


Fig. 4 Mercury cell room (cross section, schematic). (a) Basement floor; (b) floor drains; (c) cell supports with insulators; (d) supply pipes; (e) cells; (f) decomposers; (g) service walkways; (h) crane; (i) ridge ventilator; (j) ventilation air supply; (k) windows/lighting [3, p. 43].

cathode compartment that serves to impede OH^--transport (see Sect. 5.2.2). This is reenforced by a continuous flow of the NaCl electrolyte into this compartment, thus equalizing the effects of diffusion and migration of OH^--ions into the anode compartment. Figure 5 shows a cut-away view of this arrangement.

The cathodes themselves are made, for instance, of steel mesh or woven steel and the anodes are DSA in form of expanded titanium meshes. The brine (310–320 g NaCl/L) enters the anode compartment and completely covers the anodes and the cathodes (called *fingers or tubes*). The chlorine leaves the cell through an outlet in the cell head. The anolyte flows through the diaphragm into the cathode compartment because of a difference established between the anolyte level and the catholyte level. The catholyte consists of a mixture of ca. 180–210 g NaCl/L and 120–140 g NaOH/L, it is called *cell liquor*. It leaves the cell via a level control pipe on the cathode. The hydrogen produced at the same time leaves the cell through an outlet assembly on the front side of the cathodes.

A single industrial monopolar cell has geometrical dimensions of $2.5 \times 2.5 \times 2.5\ m^3$ or more. The specific load amounts to 1.5–2.5 kA m^{-2} or in total up to 150 kA, the chlorine output is about 4.5 tons per cell and day. The cell voltage is 3.0–3.5 V.

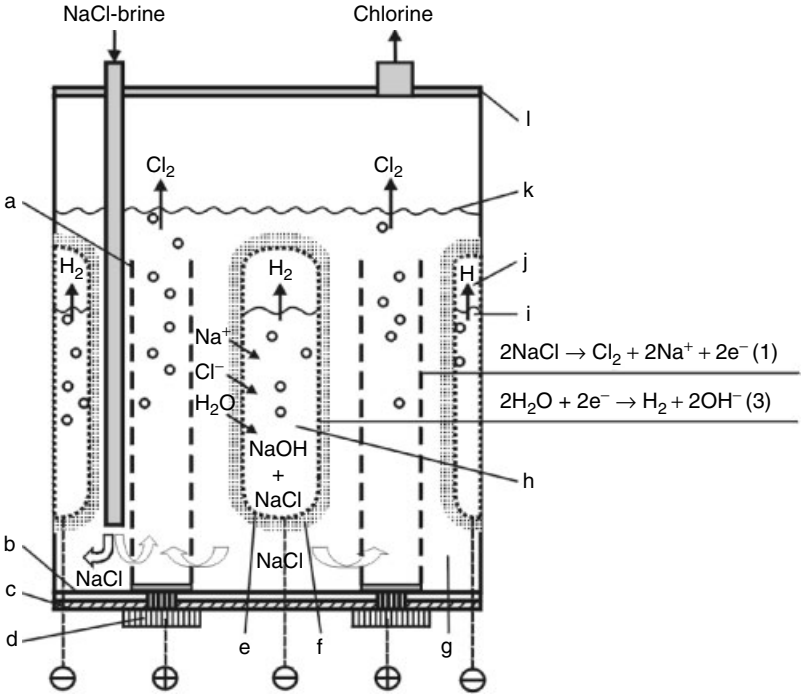


Fig. 5 Cutaway view of a diaphragm cell. (a) Anode; (b) base plate protector; (c) base plate; (d) current busbars; (e) cathode; (f) deposited diaphragm; (g) anolyte compartment; (h) catholyte compartment; (i) catholyte level; (j) hydrogen gas-room; (k) anolyte level; (l) cell head. Not included: outlet for cell liquor and hydrogen; compare text [3, p. 52].

The current efficiency reaches a value of 0.96, and the operating temperature is 90–100 °C.

Figure 6 presents a sketch of a modern monopolar diaphragm cell (there are bipolar arrangements as well).

Comparing the amalgam process and the diaphragm process beforehand, the diaphragm variety has the advantage of a lower electrolysis voltage. But the alkali product is a mixture of alkali NaOH and salt NaCl, which has to be separated (see Sect. 5.2.3.4), while the amalgam process generates pure alkali.

5.2.3.3 Membrane Process

The specific advantages of the diaphragm process (low cell voltage) and of the amalgam process (generation of pure alkali) can be combined, if in a diaphragm cell the diaphragm is replaced by an ion-exchange membrane, which is impermeable for

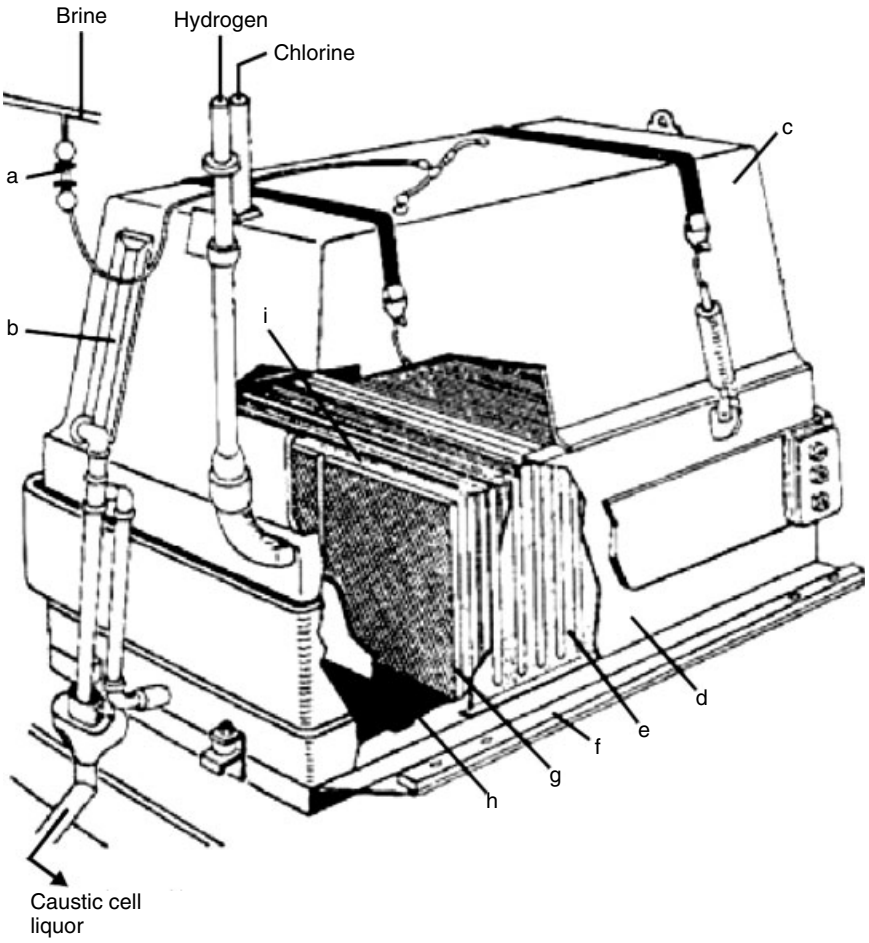


Fig. 6 ELTECH Systems MDC cells. (a) Brine feed rotometer; (b) head sight glass; (c) cell head; (d) cathode frame assembly; (e) tube sheet; (f) base plate; (g) cathode tube; (h) base plate protector; (i) DSA [3, p. 65].

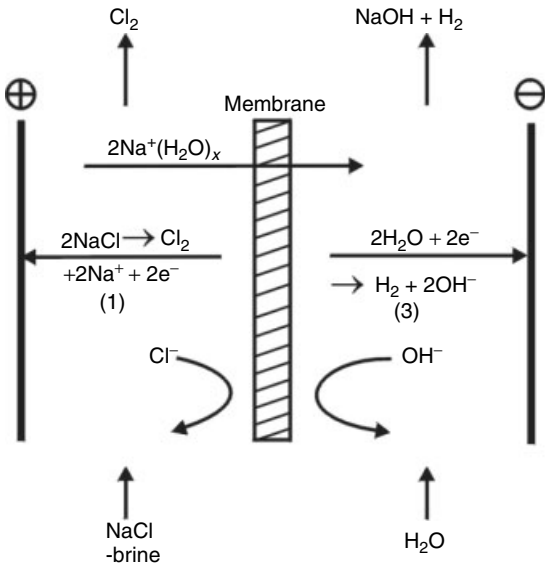


Fig. 7 Principle of the membrane process for chlor-alkali electrolysis [14, p. 441].

Cl^- and OH^--anions, but can be passed by Na^+-ions without hindrance. Figure 7 visualizes this principle, which is on the market since 1975.

The membrane is exposed to chlorine and brine on one side and strong caustic solution on the other side at temperatures of around 90 °C. Only ion-exchange membranes made of perfluoropolymer can withstand such severe conditions. The ion-exchange groups are sulfonate (SO_3^-) or carboxylate ($CHCOO^-$). Modern membranes of this kind consist of three layers, a carboxylic layer on the cathode side, a sulfonate layer on the anode side and a reinforcement layer of fabric in between. In addition, both sides are provided with a hydrophilic coating [3, p. 81; 15]. The thickness of the membrane is only one- to two-tenth of a millimeter. It must be mentioned that the hydrate water of the cations is taken along through the membrane.

Commercial membrane cells use iron or nickel (e.g. Raney-nickel) electrodes as cathodes and the anodes are from the dimensionally stable variety again. The current density goes up to 7 kA m^{-2}, the cell voltage then being 3.2 V. This is less as it would be in case of diaphragm cells at the same load, due to lower ohmic losses – the electrodes can be placed within a minimal distance near the membrane (1–3 mm, finite gap principle) or on the membrane itself (zero gap principle, compare Sect. 5.2.3.7.1).

The current efficiency amounts to ca. 95% again. The reason for losses lies in the fact that the membranes are not completely impermeable for OH^- -ions, so that reactions (6) to (8) can take place to a certain amount. Caustic alkali is generated with ca. 33 wt%, in theory it would be free of chloride-ions, but in reality Cl^--ions can penetrate the membranes, too, by a certain amount. So, the NaOH generated will contain up to 50 ppm Cl^-.

Commercial amalgam cells are strictly monopolar, and diaphragm cells exist in monopolar and bipolar variety. This holds for membrane cells, and having presented a monopolar industrial diaphragm cell as

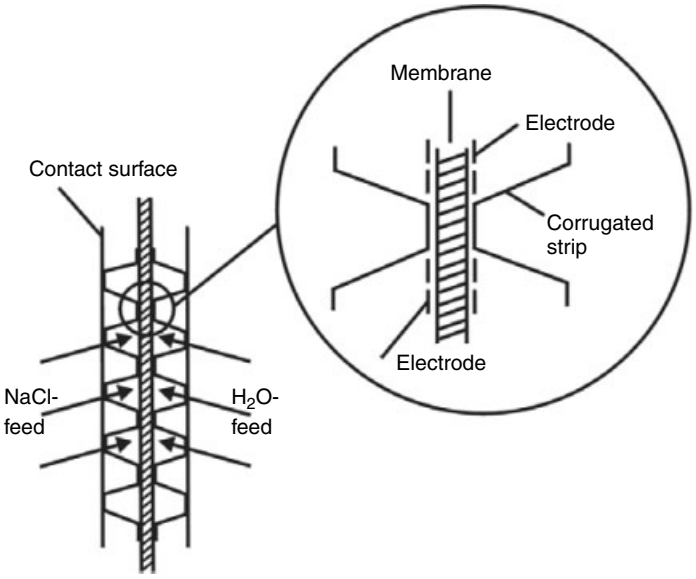


Fig. 8 Simplified cross section through a single element of a bipolar membrane cell. The electrodes are supported by corrugated bands, which are in turn supported by contact plates. Compression of the single cells together gives the series connections for the stack [14, p. 442].

an example before, a bipolar membrane cell will be presented now.

One bipolar element is shown in Fig. 8 in form of a simplified cross section.

Figure 9 shows a complete cell.

Considering a membrane surface of 2.72 m^2 for the single element and 160 elements forming the complete assembly, 81 tons of chlorine are produced a day at a current density of 5 kA m^{-2}.

Important vendors of membrane cell technology are Asahi-Kasei, Chlorine Engineers, ELTECH, and INEOS and UhdeNora.

5.2.3.4 Treatment of Educts and Products

5.2.3.4.1 **Brine** Starting point of the chlor-alkali process is usually solid NaCl in form of rock salt, solar salt or prepared from solution-mined brine. These materials of course contain impurities. In case of rock salt, the most important impurities are (in wt%) water (<3), insoluble components (<2), calcium (0.2–0.3), magnesium (0.03–0.1), sulfate (<0.8), and potassium (<0.04). But there are not only Ca^{2+}-, Mg^{2+}-, K^{+}-, and SO_4^{2-}-ions in the brine prepared from the solid salt but also Ba^{2+}- and Al^{3+}-ions and certain amounts of heavy metal ions. And last but not the least impurities generated from the process itself (ClO^{-}, ClO^{3-}) or from materials of tanks, of pipe-work and of cell components have to be taken into consideration. And of course, the solving water can be a source of impurities, too [3, p. 84].

The contaminations are harmful for the three chlor-alkali processes in different ways (as will be discussed in the following). So, the brine has to be purified, usually through precipitation by addition of carbonate and hydroxide followed by filtration steps. A more modern possibility

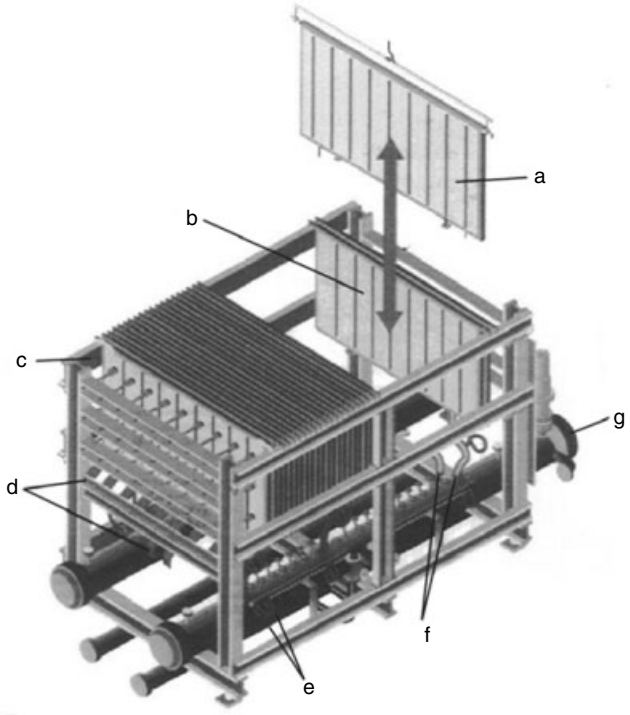


Fig. 9 Uhde BM 2.7 electrolyzer. (a) Single element; (b) contact strip; (c) cell rack; (d) busbar; (e) inlet hoses; (f) outlet hoses; (g) header [4, p. 477; 16].

is the direct use of nanofilters [17]. In any case, the incoming brine should be adjusted to a pH-value below 6 by adding HCl. This measure increases the life of the DSA-coating and neutralizes OH^--ions near the anode, thus suppressing reactions (6) to (8).

In case of mercury cells, especially heavy metal ions are precipitated galvanically at the mercury surface. In this case, parallel to the formation of sodium amalgam hydrogen gas will be evolved. The respective mechanism is discussed in detail in Sect. 5.2.3.6.1. So only traces of, for instance, vanadium, molybdenum, or cobalt are allowed in the brine (<0.01 ppm) [1].

In case of membrane cells, contaminants like Ca^{2+} (originating from the salt, maximum allowable concentration in the brine <20 ppb), Mg^{2+} (salt, <20 ppb), Sr^{2+} (salt, <4 ppb), Ba^{2+} (salt, <0.5 ppm), I^- (salt, <0.2 ppm), or Ni^{2+} (salt, tank material etc., <0.2 ppm) are examples for materials precipitating on or within the membrane, forming insoluble hydroxides, thus leading to premature damage of the membrane [3, p. 84]. To achieve these low concentrations, a secondary brine purification step by means of ion exchangers is necessary.

And last but not least it must be stated that in the case of the mercury and the membrane process the depleted brine leaving the cells must be dechlorinated before resaturation, for instance by spraying it into a vacuum of 50–60 kPa.

5.2.3.4.2 **Chlorine** Chlorine gas produced in the various electrochemical processes is saturated with water vapor at high temperatures and may also contain brine mist and traces of chlorinated hydrocarbons (originating from piping and rubber-lined vessels), and is normally at atmospheric pressure. Before the chlorine can be used it must (in this order) be cooled, purified, dried, compressed, and if necessary liquefied.

The cooling can be performed indirectly in heat exchangers (made of titanium), or directly in packed towers, where the cooling medium water is sprayed into the top and flows counter-currently to the chlorine. The cooling of course must not go below ca. 10 °C, otherwise solid chlorine hydrate will be formed, as water is still present. After cooling, water droplets and brine mist are removed by filtering elements using glass wool or porous quartz granules, alternatively electrostatic purification can take place. Organic impurities can be adsorbed by carbon filters.

The drying of the chlorine is carried out almost exclusively with sulfuric acid of 96–98 wt% in a multistage process (for instance packed towers, counter-current flow). After the drying, another demisting step, now from sulfuric acid, is necessary.

Compression is carried out with the usual mechanical means for this purpose. Care must be exercised to prevent the heat of compression from increasing the temperature enough to ignite materials in contact with the chlorine.

The quality of the compressed and dried chlorine gas from the electrolysis plant is sufficient for most applications, for example, the production of EDC for PVC. Liquefaction is necessary, if gaseous components like oxygen, carbon dioxide, nitrogen, or hydrogen have to be removed, or if chlorine is stored as a liquid. In the first case, the noncondensable gases are separated from the chlorine liquid and the latter is reevaporated. The liquefaction conditions can be selected within wide limits [3, p. 143].

Any hydrogen is concentrated within the tail gas of the liquefier and may reach the explosion limit of 4%. Therefore, the conversion of gas to liquid is limited to a yield of 90–95%. The chlorine of the tail gas is absorbed in a NaOH lye to form ''bleach'' or ''hypo,'' a sodium hypochlorite solution.

5.2.3.4.3 **Caustic Alkali** Sodium hydroxide solution from the decomposer of the amalgam process usually has a concentration of 50% and a chloride content of only 5–50 mg kg^{-1}. It is cooled down from 80–120 °C to 40–60 °C in steel pipes, Ni-, or Incolloy coolers. Any particles of graphite from the decomposer material or traces of mercury are removed by centrifuges, candle filters, or precoated leaf filters.

In the case of diaphragm cells, the cell liquor (10–12 wt% sodium hydroxide and 16–18 wt% unreacted sodium chloride) is concentrated in multieffect evaporators to 50% NaOH. The sodium chloride then is salted out and must be removed, only 1–1.5 wt% remain in the solution. Depending on customers' requirements further purification steps will follow, for example, adsorption of salt, chlorate, and carbonate by liquid ammonia or removal of metal impurities by cathodic deposition [18].

Concerning the membrane process, the caustic soda solution has a concentration of up to 32 wt%. If 50 wt% are required, evaporation can be used. The typical concentration of NaCl is 20 ppm.

If solid caustic soda is needed, further evaporation at temperatures up to 400 °C produces an anhydrous melt, from which

flakes, prills, moulded pieces, or cast blocks are formed.

5.2.3.5 Comparison of the Three Processes

This comparison starts by summing up the basic data for the three processes (Table 1).

As can be seen, the membrane process has the lowest energy costs. As the rest of the operating costs (expenses for salt, precipitants, repairs etc., costs for manpower) are roughly the same for all three processes, the membrane process has a clear advantage. This goes along with product qualities almost as good as in the case of the amalgam process and clearly better than in case of the diaphragm process (Table 2).

So, the relatively new (1975) membrane process is favored in case new plants are built. Nevertheless, most chlorine nowadays is produced now as before by the classic amalgam and diaphragm processes. But the membrane process comes up: in 1995, 52% of the total world chlorine market (41×10^6 t/a) depended on the diaphragm process, 22% on mercury cells and 21% on membrane cells, 5% others [1, p. 248]. In 2001, these data had changed to: chlorine capacity 53.5×10^6 t/a, diaphragm process 42%, amalgam process 19%, membrane process 36%, 3% others. Considering the year 2005, the share of the membrane process is estimated to be 45% [19].

Newly built membrane cell plants will cost ca. 300 000 $ per metric ton per day chlorine capacity (green-field basis). A comparison with the investment costs for the other two varieties does not make sense at the moment, because since 1991 only a few diaphragm cell plants and no mercury cell plants were built [3, p. 123].

Table 3 summarizes the advantages and disadvantages of the three chlor-alkali processes.

The last point in this comparison will be the remarks ''use of mercury, costly

Tab. 1 Basic data for chlorine production

	Mercury process	*Diaphragm process*	*Membrane process*
Decomposition voltage [V]	3.11	2.15	2.15
Theoretic specific energy usage [kWh/t][a]	2470	1705	1705
Cell voltage [V]	4	3.5	3–3.2
Current density [kA m^{-2}]	10	2.5	4–6
Electricity for electrolysis [kWh/tCl_2]	3100–3400	2300–2900	2200–2600[b]
Steam equivalent [kWh/tCl_2, 1t of steam equivalents to 400 kWh]	0	800–1000	200–400
Total energy needed [kWh/tCl_2]	3100–3400	3100–3900	2400–3000
Relative energy costs [%]	100	106	84

[a] By-product 1.13 tons of caustic soda (50%).

[b] Chlor-alkali plants in Western Europe pay depending on their capacity 5–10 Eurocents per kWh. This corresponds roughly to 20 Euros per 100 kg of elemental chlorine.

Tab. 2 Product qualities: typical compositions of chlorine, caustic, and hydrogen [3, p. 118]

Product and contents	*Process*			
	Diaphragm		*Mercury*	*Membrane*
	Unpurified	*Purified*		
Chlorine gas (from cells) [vol%]				
Cl_2	96.5–98		98–99	97–99.5
O_2	0.5–2.0		0.1–0.3	0.5–2.0
CO_2	0.1–0.3		0.2–0.5	
H_2	0.1–0.5		0.1–0.5	0.03–0.3
N_2	1.0–3.0		0.2–0.5	
NaOH solution [wt%]				
NaOH	50.0	50.0	50.0	50.0
NaCl	1.0	0.025	0.005	0.005
Na_2CO_3	0.1	0.1	0.05	0.04
Na_2SO_4	0.01	0.01	0.0005	0.0001
$NaClO_3$	0.1	0.001	0.0005	0.001
SiO_2	0.02	0.02	<0.001	0.002
CaO	0.001	0.001	0.001	0.0001
MgO	0.0015	0.0015	0.0002	0.0001
Al_2O_3	0.0005	0.0005	0.0005	0.0001
Fe	0.0007	0.0007	0.0005	0.0004
Cu	0.0002	0.0002	0.00001	0.0001
Hg	None[a]	None[a]	0.00001	None[a]
NH_3		0.001		
Hydrogen gas [vol%]				
H_2	>99.9		>99.9[b]	>99.9

[a] $<10^{-6}$%.

[b] Hydrogen gas from the mercury process contains mercury: 1–10 mg m^{-3}, depending on the purification process. The hydrogen gas from the other two processes is free of mercury.

environmental protection'' in Table 3, in the public considered as a major drawback. For a better understanding: the mercury in the electrolytic cell circulates in a closed system. All materials that come into contact with the mercury – equipment, products, auxiliary chemicals, wash water, waste gases, other waste material – may become contaminated with mercury and therefore are treated before release to the environment or are safely deposited. For the exact measurement of trace amounts and for the control of the effectiveness of the measures to reduce the emissions, analytical methods have been developed with sensitivities in the microgram region (see this encyclopaedia, Vol. 3).

Concerning the products, the decontamination from mercury has been described in Sect. 5.2.3.4.2 and 5.2.3.4.3, with exception of the hydrogen gas generated from amalgam in the decomposer. This of course has to be decontaminated, too, especially if it is used for heating or is vented to the atmosphere.

This is done firstly by cooling down the gas (2–3 °C) because the mercury content in hydrogen strongly depends on temperature (2404 mg m^{-3} at 100 °C, 2.3 mg m^{-3} at 0 °C, standard pressure).

Tab. 3 Advantages and disadvantages of the three chlor-alkali processes [3, p. 117]

Process	*Advantages*	*Disadvantages*
Diaphragm process	Use of well brine, low electrical energy consumption	Use of asbestos, high steam consumption for caustic concentration in expensive multistage evaporators, low purity caustic, low chlorine quality, cell sensitivity to pressure variations
Mercury process	50% caustic direct from cell, high-purity chlorine and hydrogen, simple brine purification	Use of mercury, use of solid salt, expensive cell operation, costly environmental protection, large floor space
Membrane process	Low total energy consumption, low capital investment, inexpensive cell operation, high-purity caustic, insensitivity to cell load variations and shutdowns, further improvements expected	Use of solid salt, high purity brine, high oxygen content in chlorine, high cost of membranes

Further, decontamination to mercury concentrations of 0.03–0.002 mg m^{-3} is possible by further cooling, by compression, by adding chlorine (formation of solid Hg_2Cl_2), by adsorption on activated carbon and by other means [3, p. 45].

The cleaning of the wastewater includes chemical removal of mercury by reducing any compounds to the metal for instance with hydrazine and subsequent filtering. Other cleaning processes, for instance using oxidation processes or ion exchange processes, are common, too.

The purification of the cell room ventilation air would require the removal of very low mercury concentrations from very large volumes of air. This would not be effective. Therefore, this concentration is kept very low by means of ''good housekeeping'' [20].

Figure 10 shows the mercury emissions from Western European chlor-alkali industry over the years from 1978 to 1996 in mg Hg/t Cl_2. With a level of less than 2 gHg/t Cl_2, the contribution of the chlor-alkali industry to the total natural and anthropogenic mercury emissions is less than 0.1%. Indeed, it may happen that a mercury plant taking the water feed from a river works as a sewage plant.

The membrane process is favored when new plants are built. But amalgam plants have economic cycle lives of 40–60 years. So, some amalgam plants will survive 20 or more future years, especially those, which produce special chemicals such as KOH, alcoholates, dithionites, and others (see Sect. 5.2.3.7.3).

In case of the diaphragm process, there are research activities now as before, for instance replacing the asbestos [22], lowering the cell voltage, using more active electrodes, and so on [23]. And there exists the possibility of a so-called retrofitting, that is, replacing the diaphragms of existing cells by membranes.

5.2.3.6 Electrode Kinetics

Introducing the electrochemical reactions taking place during chlor-alkali electrolysis in Sect. 5.2.2, it was noted that some more detailed information would follow. For better understanding, it is necessary

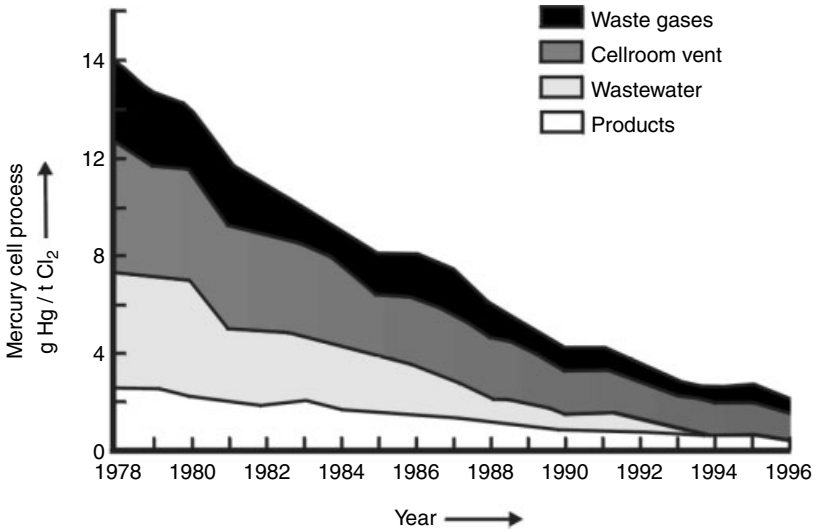


Fig. 10 Mercury emissions from the Western European chlor-alkali industry. Considering 2004 the values have been reduced to 1 g Hg/t Cl_2 [21].

that the reader has knowledge in the fields of concentration- (exactly activity-) dependence of reversible electrode potentials, of current-potential relationships under different conditions, of electrocatalysis, of activation overvoltages and of mixed potentials/mixed electrodes. This knowledge can be found in the first two volumes of this encyclopaedia, or in any textbook of electrochemistry, especially in [24].

The basic kinetics of the amalgam, the diaphragm, and the membrane cells are discussed according to Fig. 11, containing reversible electrode potentials and current-potential relationships.

5.2.3.6.1 **Amalgam Cell** Starting with the chlorine electrode, the reversible standard potential is +1.37 V (all potentials are given versus NHE), due to the sodium concentration of 6 mol L^{-1} we have to note +1.33 V. The current-overpotential curve represents an unhindered process occurring at the activated titanium material used. But we have to take into account that the reversible potential for oxygen evolution from neutral solution amounts to only +0.82 V, that is, from the thermodynamical point of view oxygen should be developed preferably. But oxygen evolution at the DSA-material is strongly hindered, the current-overpotential relationship being flat and showing an activation overpotential of more than 0.5 V. Thus in contradiction to the thermodynamical view, chlorine evolution is favored and even at current densities of 10 kA m^{-2} oxygen will contaminate chlorine only to a small extent (<2%).

The reversible potential of the sodium amalgam electrode, considering an amalgam concentration of 0.2 wt% sodium and the already mentioned sodium concentration, amounts to −1.78 V (the difference to the Na/Na^+-standard potential −2.71 V is due to the fact that in the case of amalgam the discharged sodium ions must not be incorporated into a metallic structure).

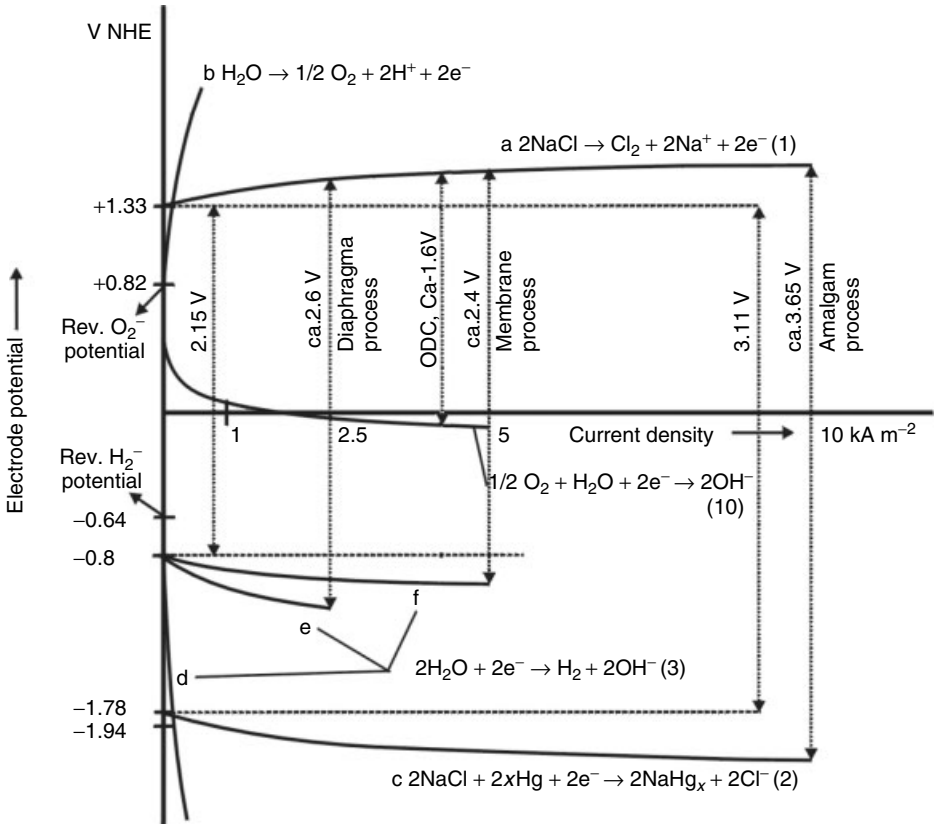


Fig. 11 Reversible electrode potentials and current density relationships (working conditions, 80 °C), ohmic losses not included. (a) Chlorine evolution at activated titanium; (b) oxygen evolution at activated titanium; (c) sodium amalgam deposition; (d) hydrogen evolution at amalgam; (e) hydrogen evolution at iron; (f) hydrogen evolution at activated nickel [4]. The acronym ODC and Eq.(10) are explained in Sect. 5.2.3.7.2.

The current-potential relationship again represents an unhindered process.

This time we have to take into account that the reversible potential for hydrogen evolution from neutral solution is far less negative (−0.4 V). But the hydrogen activation overvoltage at mercury amounts to a respectable 1.3 V. So, hydrogen evolution will start at the mercury surface at −1.7 V, that is, positive to the formation of amalgam. But according to reaction (3) then the surroundings of the surface will become alkaline and the reversible hydrogen potential is shifted to more negative values and the hydrogen evolution is drawn back. Under working conditions we have pH-values around 11 near the cathode. The reversible hydrogen potential then is −0.64 V, adding the activation overpotential of +1.3 V, so we have a beginning hydrogen evolution at −1.94 V. Again, at 10 kA m^{-2}, the contamination of the chlorine with hydrogen is tolerable.

From the reversible potentials given for the two electrodes (+1.33 resp. −1.78 V) we deduce that there is a reversible

decomposition voltage for the amalgam process of 3.11 V. The total overpotentials at the anode at 10 kA m^{-2} being 0.25 V, at the cathode 0.30 V, the electrolysis is running at ca. 3.65 V without considering ohmic losses in the electrolyte and in the electrodes, which are adding roughly an additional 0.5 V.

Coming back to the activation overpotential for hydrogen evolution: It is quite clear that if it is diminished, an explosive chlorine/hydrogen mixture will be produced. This will be the case immediately, if ions of heavy metals or other metals are precipitated at the mercury surface. So the brine must be thoroughly purified, as described in Sect. 5.2.3.4.1. This holds especially if potassium- or lithium-brine is used as a feed, because the reversible potentials of these species are more negative.

The hydrogen overpotential is diminished also, if the amalgam concentration goes up. As the reversible potential of the amalgam electrode in addition then is shifted to the negative, amalgam concentrations exceeding 0.45% become dangerous, and exceeding 0.7% the amalgam becomes highly viscous.

5.2.3.6.2 **Amalgam Decomposition** Reaction (4), representing the production of alkali and hydrogen from amalgam and water in the decomposer, corresponds to the overall reaction of a mixed electrode.

The anode reaction reads

$$2NaHg_x \longrightarrow 2Na^+ + 2xHg + 2e^- \quad (9)$$

This reaction is unhindered. The cathode reaction reads

$$2H_2O + 2e^- \longrightarrow H_2 + 2OH^- \quad (3)$$

This reaction is extremely hindered at mercury, so the amalgam is stable in aqueous solution. It is less hindered at materials like graphite, especially if activated with oxides of iron, nickel, or cobalt or by carbides of molybdenum or tungsten (the difference in the exchange current densities at mercury and at activated graphite amounts to nine orders of magnitude; the corresponding difference to reaction (9) then remains to be three orders of magnitude). So bringing the amalgam in contact with activated graphite, the decomposition starts immediately [3, p. 35].

The situation concerning the mixed electrode is given in Fig. 12. The reversible potential of the hydrogen electrode now is −0.8 V, corresponding to the strong alkaline situation. Figure 13 shows the situation at the phase boundary. Please note that the situation can also be interpreted as a short-circuited galvanic element of an open circuit voltage of one volt.

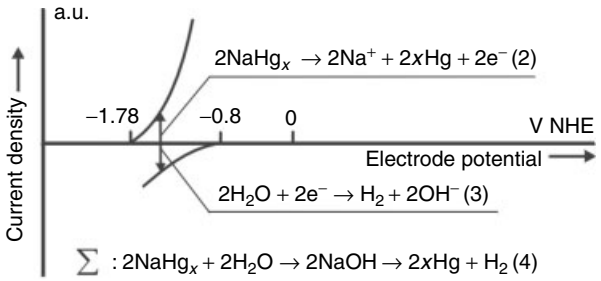


Fig. 12 The sodium/hydrogen mixed electrode.

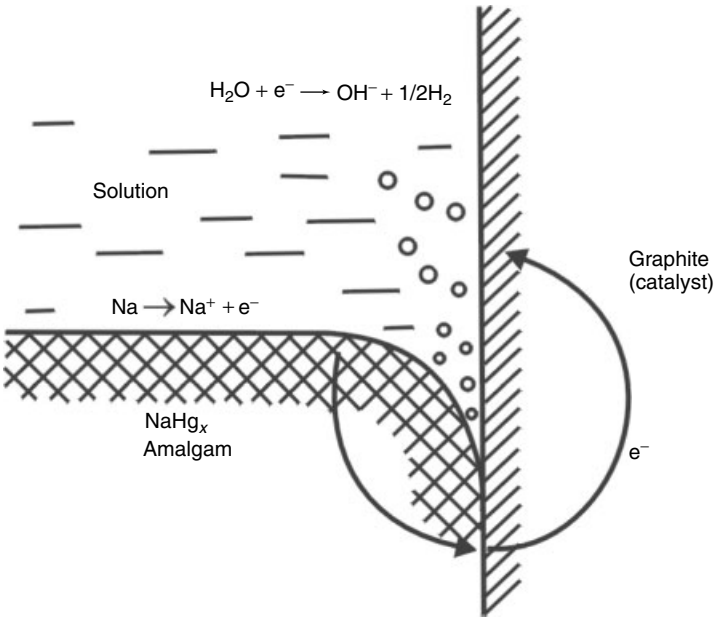


Fig. 13 Principle of amalgam decomposition.

5.2.3.6.3 **Diaphragm and Membrane Cell** Looking at the diaphragm and membrane processes (combination of Eq. (1) and (3)), the discussion concerning the chlorine anode is reproduced. At the cathode, the situation is as follows:

The reversible hydrogen potential is around −0.8 V, as the electrode works under strongly alkaline conditions. So, the reversible decomposition voltage amounts to about 2.15 V.

In case of overpotentials, we have to distinguish between the diaphragm and the membrane process, as there are different working current densities. In case of the diaphragm process ($\leq$2.5 kA m^{-2}), they amount to 0.15 V at the anode and 0.30 V at the cathode, giving 2.6 V in total (see Fig. 11). Adding the ohmic losses, the electrolysis voltage is 3.5 V (see Sect. 5.2.3.2). In case of the membrane process, due to the better conductivity of the membrane compared to the diaphragm and to the effect of activated nickel (see Sect. 5.2.3.3) the current densities increase to a value of 5–7 kA m^{-2}. The total overpotentials then are 0.20 resp. 0.15 V, which add to 2.4 V, and the total electrolysis voltage amounts to 3.0–3.2 V.

5.2.3.7 Closing Remarks

5.2.3.7.1 **Finite Gap and Zero Gap Principle** If there is a gap situated between electrode and membrane (or in other cases diaphragm), this is called *finite gap principle*. A minimum gap of roughly 1 mm is needed to ensure sufficient mass transport. Some tenth of a volt in cell voltage can be saved, if the zero gap principle is applied to chlor-alkali membrane cells. According to this principle, electrodes, which can be transmitted by the feed, are arranged directly on the surface of the membrane on both sides with adjacent transport and contact elements. This principle is used

in modern hydrochloric acid electrolysis, modern water electrolysis and in fuel cell technology, too (for more information see Sect. 5.2.4).

5.2.3.7.2 **Oxygen Depolarized Cathode (ODC) [25]** Roughly, one volt in cell voltage can be saved, if the hydrogen-evolving cathode (reversible potential in alkaline solution −0.8 V) is replaced by an oxygen reduction cathode (oxygen depolarized cathode, reversible potential in alkaline solution +0.4 V). This makes a reversible difference to the chlorine evolution (+1.33 V) of only 0.93 V. For the electrochemical reduction of oxygen we have

$$\tfrac{1}{2}O_2 + H_2O + 2e^- \longrightarrow 2OH^- \quad (10)$$

Equation (10) delivers in combination with Eq. (1)

$$2NaCl + \tfrac{1}{2}O_2 + H_2O \longrightarrow 2NaOH + Cl_2 \quad (11)$$

The lack of hydrogen formation in this process is of no importance.

The process (11) can be carried out at high current densities with the help of gas diffusion electrodes (see Chapter 8). Fig. 14 shows a schematic representation. From the current density–potential relationship of reaction (10) at an ODC, given in Fig. 11, an operating voltage around 1.6 V can be deduced (Please remember that ohmic losses are not included).

As can be seen, the caustic alkali is formed in the cathode compartment again. If air is used instead of oxygen, carbon dioxide must be removed thoroughly (see Chapter 8 of this volume).

Cells under development use setups according to the finite gap technology. Assuming a vertical ODC with the catholyte solution at the one and the gas on the other side, the hydrostatic pressure of the liquid is increasing from the top to bottom,

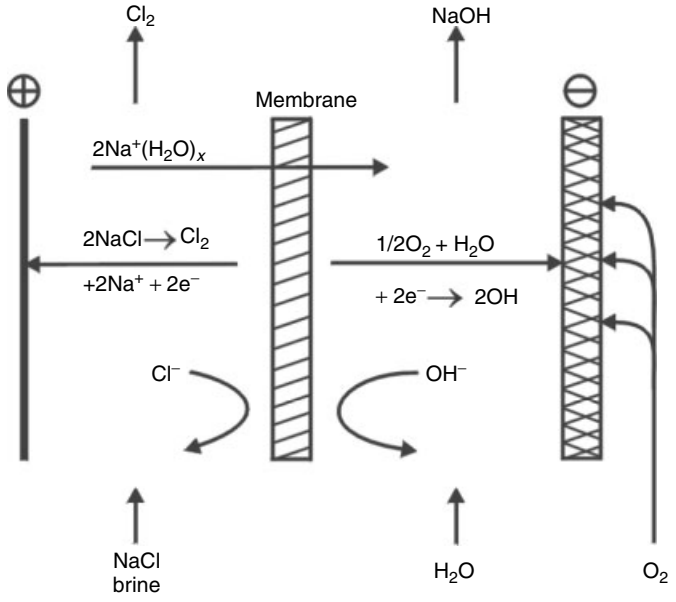


Fig. 14 Schematic representation of the chlor-alkali electrolysis using an oxygen depolarized cathode.

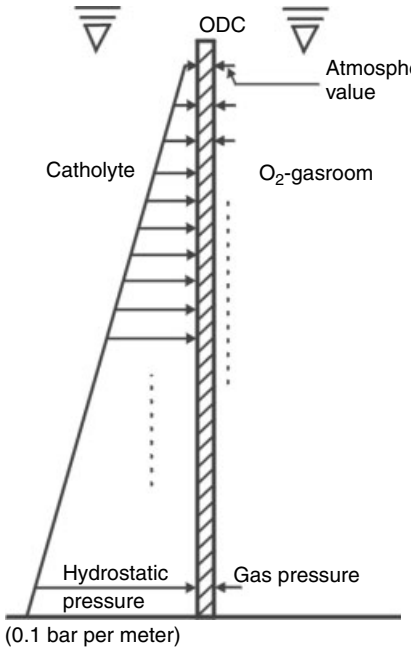


Fig. 15 Pressure situation along a vertical ODC.

while the pressure in the gas room does not depend on any coordinates. Fig. 15 demonstrates the situation considering atmospheric pressure at the top.

So, the liquid will penetrate the ODC at the bottom of the assembly, if a certain height is exceeded, depending on the permeability of the ODC. If the gas pressure on the contrary would be adjusted to the hydrostatic pressure at the bottom, then the gas would penetrate the ODC at the top and escape into the electrolytic room. In conclusion, the height of a corresponding cell is limited.

There are two developments to solve this problem, the so-called falling film principle and the gas pocket solution.

In case of the falling film principle (initiated by the Hoechst AG and continued by the UhdeNora) the central idea is to decrease the hydrostatic pressure by establishing a falling film of caustic within the gap between membrane and electrode. For satisfying results, there has to be a layer of hydrophilic material within the gap [26].

The gas pocket solution was introduced by the Bayer group. The pockets mentioned are horizontal compartments of limited height. The oxygen is fed into the lowest pocket under a pressure that equals the hydraulic pressure at the bottom. The gas then passes into the pocket above, where there is a diminished pressure, and so on (Fig. 16).

If the height of the pocket is chosen, convenient (no penetration of liquid at the bottom and no trespassing of gas in the upper region) assemblies of unlimited heights become possible [27].

Please note that the gas pockets in reality have a deepness of only some millimeters. The correct gas pressure in the consecutive pockets is ensured hydraulically without using mechanical parts [7, 14p. 445; 28].

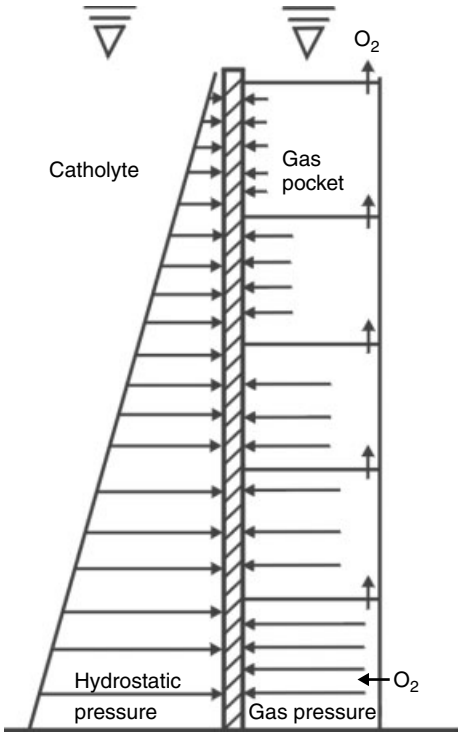


Fig. 16 Gas pocket principle.

Following the gas pocket solution, test electrolyzers of industrial size (e.g. 2.5 m^2 cross section) were equipped with ODC's and run successfully over a longer period. Current densities up to 6 kA m^{-2} are accessible, going along with cell voltages of only 2.0–2.2 V.

The catalysts used in the gas diffusion electrodes for oxygen reduction are from the Pt-group, including this metal itself or rhodium compounds. They are for instance incorporated into graphitized carbon or acetylene black. Binding material is a perfluorinated copolymer (Nafion). The conducting material necessary in the gas diffusion setup is represented by steel or nickel mesh, in some cases silver-plated for better withstanding the 32 wt% caustic alkali.

5.2.3.7.3 **Production of Alcoholates, Polyolates, and Other Chemicals from Alkali Amalgam** If the decomposer of an amalgam cell is fed with methanol instead of water, Eq. (3) will be replaced as follows

$$2CH_3OH + 2e^- \longrightarrow H_2 + 2CH_3O^- \quad (12)$$

(compare Fig. 12). In combination with Eq. (9) then we have instead of Eq. (4)

$$2NaHg_x + 2CH_3OH \longrightarrow 2CH_3ONa + H_2 + 2xHg \quad (13)$$

So, sodium methanolate is produced instead of sodium hydroxide. Potassium methanolate can be produced in this way, too.

In order to obtain a reaction rate sufficient for a technical process, the material inside the decomposer should be activated with molybdenum or tungsten carbide instead of simple iron oxide. In this case, the reaction rate using ethanol or propanol to form the corresponding alcoholates is sufficient, too. Higher alcoholates can only be produced in this way with help of a microheterogeneous electrocatalysis and/or the help of ultrasonic energy [29–31].

Nevertheless, even polyalcoholates are obtained following the reaction scheme above, if sugars protected at their anomeric center solved in aprotic media are fed into the decomposer [32].

Other products of an alkali amalgam decomposition process are sodium sulfide from polysulfide solutions, sodium dithionide from hydrogen sulfide, hydrazobenzene or aniline from nitrobenzene, and adiponitrile from acrylonitrile. Last but not least alkali metals can be prepared directly from the amalgam [33].

5.2.4 Chlorine from Hydrochloric Acid

The preparation of organic chlorocompounds usually involves chlorination according to $RH + Cl_2 \rightarrow RCl + HCl$. Then half of the chlorine necessary for the reaction ends up as either HCl gas or aqueous hydrochloric acid. One possibility for the recycling of HCl is the hydrochloric acid electrolysis to give Cl_2 and H_2 again.

$$\text{Anode}: \quad 2Cl^- \longrightarrow Cl_2 + 2e^- \quad (14)$$

$$\text{Cathode}: \quad 2H^+ + 2e^- \longrightarrow H_2 \quad (15)$$

$$2HCl \longrightarrow Cl_2 + H_2 \quad (16)$$

5×10^6 tons of chlorine are produced in this way per year, which amounts to 10% of the chlorine produced by the chlor-alkali process.

Conventional hydrochloric acid electrolyzers consist of 30–36 individual cells connected in series (bipolar arrangement). The cells are formed from vertical electrode plates manufactured from graphite, between which there are diaphragms (for instance made from PVC fabric, distance to the plates ca. 6 mm). The feed with hydrochloric acid (22 wt%, identical for anode and cathode compartment) and the removal of the gases produced take place according to the filter press principle (see Chapter 2). Chlorine leaves the cell with the anolyte, hydrogen with the catholyte.

The reversible potential of the hydrogen evolution is around 0 V (remember: all potential values are given in V vs. NHE in this chapter and now we have acidic conditions). For the chlorine electrode it is ca. 1.33 V again, so we have a theoretical cell voltage of 1.33 V, the operating value being around 1.9 V at 4.8 kA m^{-2}. The theoretical specific energy usage per tonne of chlorine then amounts to 1005 kWh, and to 1450 kWh in reality.

The last value can be lowered to 1300 kWh t^{-1}, if the diaphragms are replaced by cation exchange membranes of the sulfonate type pretreated to conduct hydronium ions [3, p. 133].

Oxygen depolarized cathodes can be used in hydrochloric acid electrolysis, too.

$$\text{Anode}: \quad 2HCl \longrightarrow Cl_2 + 2H^+ + 2e^- \quad (17)$$

$$\text{Cathode}: \quad \tfrac{1}{2}O_2 + 2H^+ + 2e^- \longrightarrow H_2O \quad (18)$$

$$2HCl + \tfrac{1}{2}O_2 \longrightarrow Cl_2 + H_2O \quad (19)$$

Reaction (18) represents the oxygen reduction process in acidic solution, the reversible potential then being +1.23 V. This leads to a reversible cell voltage of 0.2 V. The theoretical energy usage per tonne of chlorine in this way is unbeatable 160 kWh.

The ODC in this case may form together with the membrane an unit, that is, the porous electrode is (including the current distributor) directly attached to the membrane surface. The electrocatalyst needed for oxygen reduction in this case is located at the membrane/electrode boundary. Concerning the anode, the finite gap principle may be used. Figure 17 gives a scheme of the principle.

The setup becomes possible because the cathodic reaction product (water) can easily be drained through the porous ODC without affecting the water content of the membrane. And there is no pressure compensation needed due to the fact that the ODC is not affected directly by the hydrostatic pressure.

Producing chlorine from hydrochloric acid using ODCs in the described

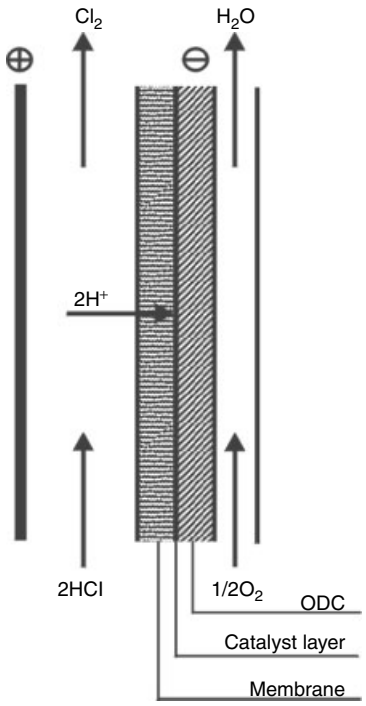


Fig. 17 Scheme of producing chlorine from hydrochloric acid using an Oxygen Depolarized Cathode (ODC). The ODC/catalyst/membrane unit is displayed in more detail in the coming Figure 23.

way – and using rhodium compounds as electrocatalyst – has been pushed by the Bayer Group. The success could be proven up to full size (2.5 m^2) cell elements. A 20 000 tons per year plant was started in 2003. Real cell voltages of 1.1 V at 2 kA m^{-2} and of 1.35 V at 4 kA m^{-2} are reported. So chlorine from hydrochloric acid can be produced with an energy usage of only 870–1070 kWh t^{-1} [7, 14, p. 447; 34].

Another development, but not yet being introduced into practice, is the electrolysis of anhydrous hydrochloric acid [3, p.135]. Again, a cation exchange membrane is used to transfer protons from the anode to the cathode side. The development aims at the production of dry chlorine directly – in all other cases of hydrochloric acid there are water contents, which have to be removed (see Sect. 5.2.3.4.2).

The subject of Sect. 5.2.1 to 5.2.4 was the electrochemical production of chlorine and of caustic alkali. It shall be noted in addition, that chlorine is obtained as a by product if metal chlorides like NaCl, $MgCl_2$, or others are electrolyzed in their molten form to produce the metals (see Chapter 4). From aqueous NaCl, caustic alkali and hydrochloric acid can be produced without generation of chlorine, using an electrodialytic separation process [35].

5.2.5 Chlorine–Oxygen Compounds [1, 4, 9, 10]

The basic electrochemical reactions ruling the generation of chlorine, caustic alkali, and chlorine–oxygen compounds were introduced in Sect. 5.2.2. For better reading they shall be repeated here

$$6NaCl \longrightarrow 3Cl_2 + 6Na^+ + 6e^- \quad (1)$$

$$2H_2O + 2e^- \longrightarrow H_2 + 2OH^- \quad (3)$$

$$3Cl_2 + 6OH^- \longrightarrow 3ClO^- + 3Cl^- + 3H_2O \quad (6)$$

$$3ClO^- + \frac{3}{2}H_2O \longrightarrow ClO_3^- + 2Cl^- + 3H^+ + \frac{3}{4}O_2 + 3e^- \quad (7)$$

$$ClO_3^- + H_2O \longrightarrow ClO_4^- + 2H^+ + 2e^- \quad (8)$$

The following text concentrates on Eqs. (6) to (8) and a few more electrochemical and chemical processes. Concerning

the spacious variations of using the chlorine oxygen compounds for bleaching, disinfection or water treatment, information on these items is only given in those few cases, where it is needed for better understanding.

In case of electrolytic production of chlorine oxygen compounds, questions like purification of brine and of products arise, too. These questions are solved in analogy to the means described in Sect. 5.2.3.4 and shall not be repeated here.

5.2.5.1 **Hypochlorite**

According to Eqs. (1), (3), and (6), hypochlorite will be formed in undivided electrolysis cells fed with sodium chloride. Combining the equations mentioned, we find as overall reaction

$$3NaCl + 3H_2O \longrightarrow 3H_2 + 3NaClO \quad (20)$$

Anode and cathode materials are equivalent to those used in the diaphragm process. The pH-value of the cell liquor is adjusted from 10 to 12. Low temperatures near ambient values suppress the disturbing reaction (7).

At the cathode of a hypochlorite cell this species can be reduced to chloride

$$ClO^- + H_2O + 2e^- \longrightarrow Cl^- + 2OH^- \quad (21)$$

So usually only diluted aqueous ClO^- is produced in the described way (10 g L^{-1} of NaClO). The process can be carried out even with simple seawater. This possibility for instance is used in case of disinfection of the cooling water needed in coastal power plants. In a by-pass a small electrolysis cell is installed. The cell then may be constructed as a tube containing the electrodes.

If higher concentrations are needed, chlorine and caustic alkali produced for instance in membrane cells are reacted outside the cell to form the so called ''bleach'' or ''hypo,'' sodium hypochlorite, according to Eq. (6).

Commercial sodium hypochlorite solutions contain usually 170–220 g L^{-1} (14–15 wt%) ''available chlorine,'' and it can be transported in vessels made of PVC, PE, or PP. Of course, chlorine and caustic alkali can be transported instead to the consumer and made to react there. But handling chlorine or bleach is dangerous. For this reason it is becoming usual to prepare the bleach at the place of use from brine directly. Especially, if small quantities and/or diluted preparations are necessary, this way is favored.

5.2.5.2 **Chlorite**

Chlorites are performed chemically from chlorine dioxide, caustic alkali, and hydrogen peroxide

$$ClO_2 + 2NaOH + H_2O_2 \longrightarrow 2NaClO_2 + O_2 + H_2O \quad (22)$$

ClO_2 is made from $NaClO_3$ and HCl (see Sect. 5.2.5.5). $NaClO_3$ is prepared electrochemically from NaCl (Sect. 5.2.5.3). H_2O_2 can be prepared electrochemically, too (Sect. 5.6). It is needed in reaction (22) as a mild reducing agent for hindrance of disproportionation of the ClO_2 to give chloride and chlorate.

So, we have, summing up the processes, the sequence $NaCl \rightarrow NaClO_3 \rightarrow ClO_2 \rightarrow NaClO_2$.

This is somewhat complicated and expensive. Nevertheless, the usage of chlorite in the bleaching of textiles and the treatment of water and of wastewater is increasing due to certain advantages (e.g. fiber-protecting mild oxidant, safer handling of the chemicals, no decomposition

during storage, suppressed formation of tri-halomethanes).

5.2.5.3 Chlorate

On the basis of reactions (1), (6), and (7) we can see that chlorate can be produced in an undivided cell in a consecutive process consisting of an electrochemical, a chemical, and another electrochemical process, needing nine electrons to generate a chlorate anion. But there is a better pathway, using the hydrolysis of elemental chlorine in water, depending on the pH-value

$$Cl_2 + H_2O \longrightarrow HOCl + H^+ + Cl^- \quad (23)$$

The acid formed undergoes dissociation

$$HOCl \longrightarrow ClO^- + H^+ \quad (24)$$

and ClO_3^- is formed according to

$$2HOCl + ClO^- \longrightarrow ClO_3^- + 2Cl^- + 2H^+ \quad (25)$$

Summing up reactions (1), (3) and – under adjustment of the stoichiometric factors – reactions (23) to (25) we get for the overall reaction

$$NaCl + 3H_2O \longrightarrow NaClO_3 + 3H_2 \quad (26)$$

Now there is only one electrochemical step included (Eq. (1)) and instead of nine electrons only six are needed for the generation of a chlorate anion. The lack of oxygen generated in the process is of no importance.

So considering a technical process it is favorable for electric energy usage to suppress the electrochemical oxidation of the ClO^- (reaction (7)) and to enhance the reactions (23) to (25), that is, the chemical reaction sequence.

First, the hydrolysis reaction must be supported by adjusting the pH-value from 6 to 7. The ratio between the homogeneous chemical reaction sequence forming the ClO^- and the heterogeneous electrochemical reaction sequence forming ClO^- is enhanced by providing sufficient reaction volume compared with the electrode surface. This can be done by arranging the undivided electrode system within a flow system, which carries the resulting liquor into a reaction room. There is a second advantage: the ClO^- generated is transported away quickly from the oxidizing anode.

The said ratio is enhanced, too, by using higher temperatures, for instance, near the boiling point. The reason is that chemical reactions usually are much more accelerated by temperature compared with electrochemical reactions (the activation energy of an electrochemical reaction is lowered by the overpotential applied, thus lowering the temperature dependence of the rate constant [24, p. 166]). In the case of a flow system, that is, different locations for the electrolytic process and the chemical sequence, the first location is cooled by cooling the inlet brine, while the higher temperature needed for a quicker chemical reaction is produced by the ohmic losses within the cells and the heat of the reactions.

Further improvements are given by a high concentration of NaCl (possible reason: displacement of water molecules needed for reaction (7) from the anodic reaction site).

Of course, losses in current efficiency can originate from the cathodic reduction of ClO^- (see Sect. 5.2.5.1) and of ClO_3^- already generated in the electrode compartment

$$ClO_3^- + 3H_2O + 6e^- \longrightarrow Cl^- + 6OH^- \quad (27)$$

Both reduction processes are inhibited by addition of chromate (3–6 g L^{-1} sodium chromate) to the cell liquor.

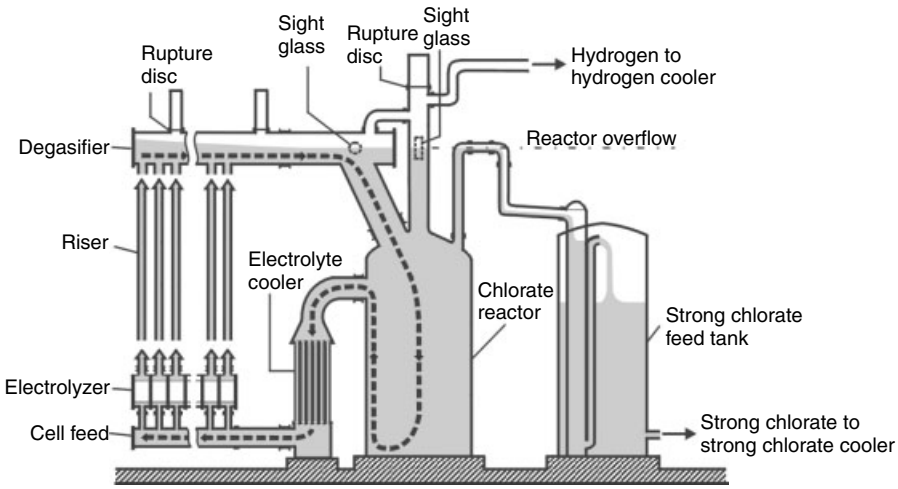


Fig. 18 Example for a technical chlorate cell (Kvaerner Chemetics) [4, p. 532; 36].

Fig. 18 gives an example for a technical chlorate cell.

Cooled and saturated brine is fed into the electrode compartments, where electrolysis takes place between steel cathodes and DSAs separated by a 3–5-mm cleft at current densities of 5 kA m^{-2} and at a cell voltage of 2.9–3.3 V.

The generated hydrogen gas forms a gas-lift pump effect, providing the flow mentioned before. Then the hydrogen is separated from the liquor flowing into the reaction compartment. Cooling, pH-adjustment via addition of HCl and addition of chromate close the liquor circuit.

Calculating the theoretical energy use of the process, we have to remember the electrolysis of aqueous NaCl at a pH-value of 6–7 as a starting point. So there exists a theoretical decomposition voltage of 1.74 V, the reversible potential of the hydrogen electrode (3) then being −0.41 V compared to 1.36 V for the chlorine electrode. So the theoretical energy usage per ton of chlorate amounts to 2600 kWh. (Remember the usage of six electrons!). In reality, considering a current efficiency of 95%, we have a value of 4600–5300 kWh.

Latest developments in the field snatch up and rename an older setup: electrolysis block and cooling are integrated into the bottom of the chemical reaction vessel (Single Vessel Technology) [4, p. 533; 37].

5.2.5.4 Perchlorates

The technical production of perchlorates follows the anodic oxidation of chlorate presented in Eq. (8). That means that chlorate (400–700 g $NaClO_3$/L) is fed into an undivided cell at a pH-value of 6–10. The cathode reaction is again Eq. (3). The overall reaction reads accordingly

$$ClO_3^- + H_2O \longrightarrow ClO_4^- + H_2 \qquad (28)$$

The reversible standard potential of the anodic reaction (8) is 1.19 V (please remember that all potentials are noted versus NHE). This is in the near neighborhood to water oxidation (1.23 V). So anode materials possessing high overpotentials for oxygen evolution have to be used, for instance platinum, platinum coated materials like titanium, or lead dioxide.

The use of less expensive (PbO_2) anode materials results in less overpotentials for oxygen evolution, that is, less current efficiencies. Nevertheless, this can be more economically compared with using platinum materials, as the loss of platinum from the electrodes can amount to 7 g per ton perchlorate generated [10].

Cathodes are made from different steel materials; electrolyte additive is sodium chromate like in the case of the chlorate electrolysis. This cannot be done in the case of lead dioxide anodes, because the oxygen evolution would be catalyzed in this way. The operation temperature is 35–50 °C.

Technical perchlorate cells are operated at up to 5 kA m^{-2}, the cell voltage then being up to 6.5 V (due to high ohmic losses in the cell) leading to a specific energy use of 2500–3000 kWh per ton $NaClO_4$. But it must be noted that the corresponding values for the educt sodium chlorate should be added [38].

Further it must be noted that perchloric acid, normally generated chemically from saturated sodium perchlorate solution and hydrochloric acid, can be generated electrochemically by chlorine oxidation, too [39].

5.2.5.5 Chlorine Dioxide

As mentioned in Sect. 5.2.5.2, chlorine dioxide is produced from electrochemically generated chlorate. In more detail, chlorine dioxide is made chemically from ClO_3^- by reaction in acid solution with a reducing agent. This agent depends on the local possibilities. One example is hydrochloric acid:

$$NaClO_3 + 2HCl \longrightarrow ClO_2 + \tfrac{1}{2}Cl_2 + NaCl + H_2O \quad (29)$$

But ClO_2 can be produced by electrolysis, too. In this case, sodium chlorite solution is introduced into the anode compartment of a divided cell and oxidized anodically [4, p. 535].

5.2.6 Additional Aspects

The present days are characterized by a temporary imbalance in the demand for alkali and for chlorine. If there is no economical way for getting rid of the chlorine surplus, it could be favorable to produce alkali without an accompanying production of chlorine.

This can be done using sodium sulfate instead of sodium chloride as the source of alkali. Na_2SO_4 is found dissolved in the wastewater of numerous chemical processes and has its problems concerning disposal, too (highly water soluble, corrosive concerning concrete).

Feeding aqueous Na_2SO_4 for instance into the anode compartment of a membrane cell (see Figs. 7 to 9), oxygen will be produced at a suitable anode material according to

$$2H_2O \longrightarrow O_2 + 4H^+ + 4e^- \quad (30)$$

Sodium ions are transferred into the cathode compartment via the membrane and form the desired sodium hydroxide as before. The anolyte leaves the cell in the form of a mixture of Na_2SO_4 and sulfuric acid.

But the conventional membrane setup described is only one possibility of a sodium sulfate electrolysis. Other proposals include the use of an ODC or even an HDA (Hydrogen Depolarized Anode) for cell voltage reduction or the use of anion exchange membranes instead of the cation variety [7]. Last but not least electrodialysis may be used: in a three compartment unit the salt is converted into free acid and the free base [40].

In the case of producing alcoholates or other chemicals from alkali amalgam (see

Sect. 5.2.3.7.3), in the future it may be favorable to produce the amalgam without generating chlorine in the mercury cell. A possible way is described in [41].

5.3 Fluorine [42, 43]

5.3.1 Introduction

The world's production of elemental fluorine amounted to around 1990 to 17 000 tons per year; it is concentrated in North America, Europe, and Japan. This production is exclusively electrochemical, because F_2 is the most highly oxidizing stable species known and so it cannot be industrially obtained by the action of another oxidant. To be exact: there are chemical reactions, which produce elemental fluorine, like $K_2NiF_6 + KF \rightarrow K_3NiF_6 + 1/2F_2$ (at 673 K, [44]). But then it has to be taken into account that elemental fluorine was needed to prepare the fluorine complex of Ni(IV). And the preparation method described in [45] is only of laboratory scale.

Of utmost technical importance is the production of UF_6, a necessity that generated the fluorine industry 65 years ago. Of technical importance too is the production of strong (e.g. ClF_3) or mild (IF_6) fluorinating agents in general or of fluorinated species in particular, for example, fluorinated oils and greases, perfluorosulfonic acid or refrigerants. Further applications are the production of graphite fluorides CF_x and of volatile metal hexafluorides used in chemical vapor deposition.

5.3.2 Electrochemical Reactions

In tables concerning reversible standard electrode potentials a value of 2.85 V NHE is found for the aqueous $F_2/2F^-$-half cell, the second half cell being $H_2/2H^+$. The decomposition voltage is 2.85 V for aqueous HF, too. Of course, this value is impossible to achieve, not to mention the higher voltages necessary for the actual decomposition process. So it turned out that the only method for production of fluorine is the electrolysis of anhydrous hydrogen fluoride containing dissolved potassium fluoride, which produces the needed conductivity (in the range of 0.1–0.2 $Ohm^{-1}cm^{-1}$). The cell reaction then is

$$2HF \longrightarrow H_2 + F_2 \qquad (31)$$

Under the conditions achieved in industrial electrolyzers of today (KF·2HF media, 85 °C), the decomposition voltage amounts to 2.9 V. Sufficient stable electrode materials are carbon (anode) and copper or steel (cathode), steel being the cheaper choice.

Along with the anode reaction, the so-called anode effect, a phenomenon often observed in fused salt electrolysis (see Chapter 4), may occur. In the present case, it may be due to a surface film of the type CF_x formed on the anode material. This film on the one hand protects the carbon against destruction (and is the reason for high anodic overpotentials) in normal operation and, on the other hand, under more or less known conditions may block electron transfer completely. These conditions depend strongly on the electrolyte composition (purity) [46, 47]. Additives, such as lithium fluoride, may be helpful in preventing the anode effect by wetting the electrode material.

5.3.3 Technical Process

Figure 19 shows an example for a classical fluorine cell developed for the French

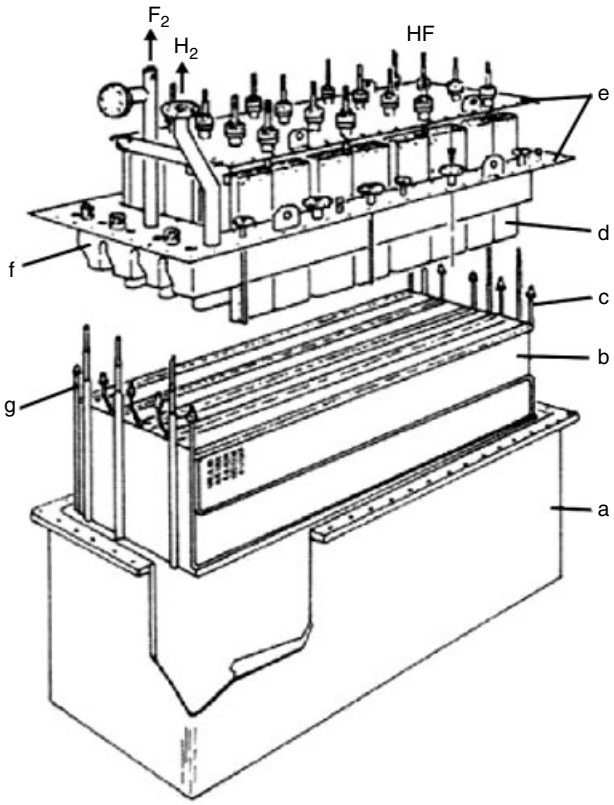


Fig. 19 Classical fluorine cell (Comurhex – Atochem, France). (a) Cell tank; (b) cathode; (c) cooling water inlet; (d) anode; (e) cell covers; (f) gas separation skirt; (g) cooling water outlet.

atomic program. The cell tank is made from Monel (a nickel-copper alloy showing better resistance against corrosion than steel, which, by the way, can also be used). Cathodes are made from steel, anodes from carbon. Monel, magnesium, or steel skirts are used for the separation of the gases. This setup shows a current efficiency of 90–95%. Other constructions use ''diaphragms'' (i.e. extended skirts with a perforated immersed part). These give higher current efficiencies and prevent broken anodes from touching cathodes. The fluorine leaving the cell contains 7–10% HF, the same holds for the hydrogen. HF is extracted from the gases and re-used. Other impurities are dealt with in the last paragraph of this section.

The usual operating voltage of a fluorine cell is 8.5–10.5 V at 1–2 kA m^{-2}. The high voltage is due to – in comparison to other electrolytic processes – relatively low electrolyte conductivity, long current paths, and especially due to the high anodic overpotential mentioned. The specific energy consumption so amounts to about 14 000 kWh/t F_2. This is about four times the theoretical value.

The large energy loss results in the necessity to remove a large quantity of heat

from the operated cell. This can be done by a water jacket, electrolyte circulation, or Monel cooling tubes that are suspended through the cell cover. When the cell is shut down, these means can be used for keeping the electrolyte molten (KF·2HF has a melting point of 71.7 °C).

Incidentally, the handling of the melt of hydrogen fluoride and of course fluorine itself is somewhat problematic, but the precautionary measures are well known (for instance, grease must be kept out of the system, because otherwise highly exothermic combustions will take place, melting or bursting piping or valves).

It should be noted, that water contents in the melt leads to the formation of O_2, CO_2 (carbon anodes!) and OF_2 in the anode compartment. In this connection the electrolyte can be purified with fluorine or by so-called pre-electrolysis: in a separate cell (with nickel–anodes) water is decomposed using a suitable low cell voltage. Further purification can take place in the production cell itself by using for several hours at the beginning of the process a reduced current density (ca. 0.2–0.3 kA m^{-2}, value estimated). The resulting fluorine contaminated with O_2, CO_2 and OF_2 is rejected. Unavoidable is the formation of fluorinated carbon, such as CF_4 in traces (ca. 30 ppm).

5.4 Hydrogen, Oxygen, and Ozone [7, 48–51]

5.4.1 Introduction

In 1989, $500 \times 10^9 m_N^3$ or 43×10^6 tons of hydrogen were produced worldwide, only

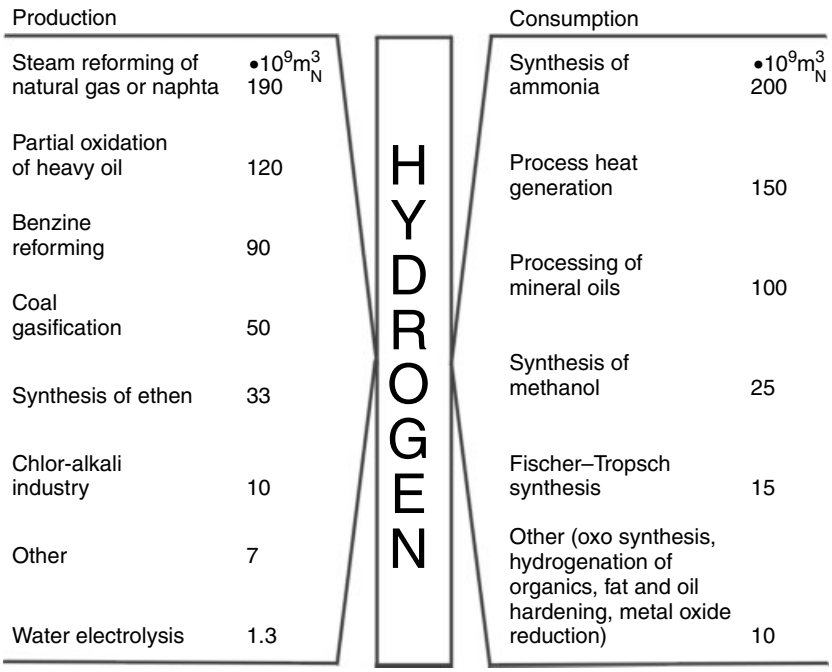


Fig. 20 Production and consumption of hydrogen [52, p. 464; 53].

0.25% of this amount by water electrolysis. Figure 20 gives a summary of production and consumption.

The minor importance of water electrolysis is due to the fact that the electrolytic process is only competitive when the cost of the electricity is low. Only in the neighborhood of hydroelectric power stations, hydrogen and subsequently ammonia and other nitrogen compounds can be prepared economically via electrolysis. This is the case, for instance, in Egypt (Assuan), Norway (Ryukan), and in Canada, where electrolysis plants up to 35 000 m_N^3/h are located. Nevertheless, there is a market for electrolyzers producing 3–100m_N^3/h hydrogen, too. These are used in case there is a limited local demand for the gas, and when the transport from the manufacturer and the storage would not be economical.

However, the electrolysis of water has received considerable recent attention in the context of the so-called ''hydrogen economy,'' in which hydrogen gas is employed as an energy vector.

Hydrogen is easily transported and stored in pipes, and in energy terms, its transport is a factor of 5 cheaper than the costs of transporting electricity through high-voltage cables. Looking further ahead, the exhaustion of fossil fuels will offer the opportunity of a complete transition to the ''hydrogen economy''. In this economy hydrogen gas, generated by nuclear power (including fusion reactions) or by solar or wind power–operated water electrolyzers is transported by pipes to consumers. There it will be used to generate power in fuel cells, as an industrial raw material, as a source of process heat through combustion, as an energy vector for transport systems or to provide energy for the home.

As regards the oxygen, the situation is similar. Along with $1.3 \times 10^9 m_N^3$ of hydrogen $0.65 \times 10^9 m_N^3$ of pure oxygen are produced worldwide by electrolysis. Against that, in the United States of America, $13 \times 10^9 m_N^3$ of so-called ''commercial oxygen'' was produced in 1989 from air by cryogenic separation, pressure swing adsorption or membrane separation. Commercial oxygen is one of the world's most significant industrial chemicals, used for instance in metallurgy, production of chemicals, bleaching, and wastewater treatment.

5.4.2 Electrochemical Reactions

As will be described in Sect. 5.4.2, there are three possibilities for the technical process, differing in the composition of the electrolyte. Alkaline, acidic, and ceramic systems are used (the ceramic system employs the so-called stabilized zirconia, which serves as an O^{2-}-ion conductor). So, we have to distinguish between three possibilities for realizing the cell reaction $H_2O \rightarrow H_2 + 1/2O_2$, too (Table 4).

It can be derived directly from the cell reaction that the reversible open circuit

Tab. 4 Possibilities for realizing the cell reaction $H_2O \rightarrow H_2 + 1/2O_2$

System	*Cathode*		*Anode*	
Alkaline	$2OH^- \rightarrow 1/2O_2 + H_2O + 2e^-$	(10)	$2H_2O + 2e^- \rightarrow H_2 + 2OH^-$	(3)
Acidic	$3H_2O \rightarrow 1/2O_2 + 2H_3O^+ + 2e^-$	(18)	$2H_3O^+ + 2e^- \rightarrow H_2 + 2H_2O$	(15)
Ion-conducting	$O^{2-} \rightarrow 1/2O_2 + 2e^-$	(32)	$H_2O + 2e^- \rightarrow H_2 + O^{2-}$	(33)

voltage does not depend on the pH-value. Under standard conditions, it is 1.23 V.

Although the reactions look quite simple, they follow complicated sequences. This has been dealt with in this encyclopaedia Vol. 7 and shall not be repeated here. Additional information will be found in [14, pp. 375–383; 24, pp. 281–283].

Instead of oxygen, ozone may be produced anodically in the acidic system

$$H_2O \longrightarrow 2H^+ + \frac{1}{3}O_3 + 2e^- \qquad (34)$$

The cell reaction then being

$$H_2O \longrightarrow \frac{1}{3}O_3 + H_2 \qquad (35)$$

with a decomposition voltage of 1.46 V. So, under normal conditions the reaction $H_2O \rightarrow H_2 + 1/2O_2$ will be preferred.

5.4.3 Technical Processes

5.4.3.1 Classical Water Electrolysis

The major part of water electrolyzers operated today use alkaline electrolytes and follow the bipolar design. Figure 21 represents the principle setup.

In order to minimize ohmic losses, 6–8 N potassium hydroxide is used as conducting agent at 80 °C. Owing to a temperature coefficient of $dE_D/dT = -0.85 mV K^{-1}$, the reversible decomposition voltage E_D then amounts to 1.19 V. The electrodes originally were situated close (≥1 mm, finite gap technology) to the diaphragm soaked with electrolyte, the diaphragm being necessary for gas separation and consisting of asbestos (conventional cells). Nowadays, the zero gap technology is used more and more, so the electrodes are made from metal wire mesh pressed onto the diaphragm. The latter is, for instance, made from ceramics (porous NiO) or from polymers (polysulfones). The electrodes being made from nickel or steel before, now use materials such as activated nickel (e.g. Raney-nickel) or nickel alloys (advanced classical cells) [54].

In all cases, the electrode/diaphragm/electrode-design must ensure that gas bubbles that form are removed rapidly from the electrolyte to avoid increases in cell resistance [53].

Figure 22 demonstrates the achievements resulting from the transition from conventional to advanced cells.

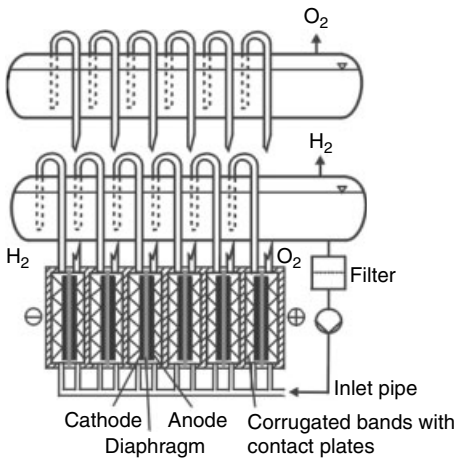


Fig. 21 Principal setup of a classical water electrolyzer.

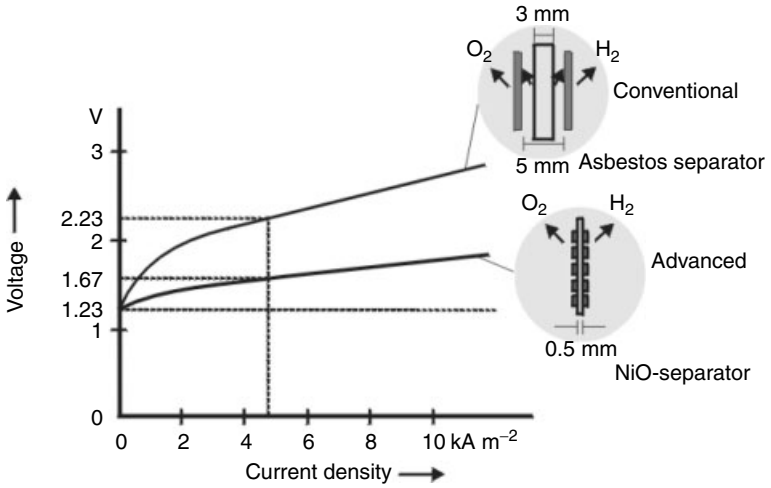


Fig. 22 Current density – voltage diagrams of a conventional and an advanced water electrolyzer [52, p. 467].

As can be seen from Fig. 22, using the advanced cell design, a cell voltage of only 1.67 V at about 4 kA m^{-2} becomes possible, the energy requirement per m^3_N then being 4 kWh. Electrolyzers up to a load of 1 MW are commercially available [53].

Fig. 22 in comparison to Fig. 11 clearly shows that in the electrolyte as well as in the water feed, chloride anions cannot be tolerated (contents must be less than 25 mg L^{-1}), owing to interference at the anode by chlorine evolution. So pure deionized water (conductivity $< 5 \times 10^{-6}\ \Omega^{-1}cm^{-1}$) is used [38].

Outside the cells, the dried gas is pressurized by multistage reciprocating or screw compressors. But the pressure can also be generated in the electrolyzer itself, remembering that after Nernst's equation the decomposition voltage for the reaction $H_2O \rightarrow H_2 + 1/2O_2$ amounts to $E_D = 1.23 + 0.045 \lg p$ (at standard temperature). This means an additional cell voltage of 45 mV per decade of pressure is needed. By the way, if within the electrolyzer a pressure is built up, the gas bubbles will shrink, lessening the ohmic losses in this way. This effect will compensate the additional demand of cell voltage to a certain extent. Another compensation is given by the fact that under pressure the temperature of operation can be augmented to 120 °C, saving some other millivolts. Realization of gas pressures exceeding ca. 200 atmospheres in the cells is nevertheless problematic, as there are problems of pressure regulation, and the higher costs for the pressure tight cell have to be compared to the costs of compressors.

5.4.3.2 Membrane Electrolyzers

Membrane electrolyzers for water decomposition use cation exchange membranes of the type used in the chlor-alkali membrane process, in this case made H^+-ion conducting by different pretreatment. So, the ions generated at the anode ($H_2O \rightarrow 2H^+ + 1/2O_2 + 2e^-$) pass through the membrane and form at the cathode hydrogen ($2H^+ + 2e^- \rightarrow H_2$).

In addition to a – compared with a diaphragm – better conductivity of the membrane, advanced electrocatalysis is applied for cell voltage reduction: electrocatalyst particles (for instance platinum or Ni–Zn alloy) are deposited in the boundary layer adjacent to a porous electrode material [55]. Figure 23 shows the principle.

100 kW electrolyzers have been built commercially, showing a cell voltage of only 1.55 V at a load of 5 kA m^{-2}. An additional advantage is a reduced volume of the unit compared to the classical variety.

But membrane electrolysis suffers a severe drawback: the high costs for the solid polymer electrolyte. If membranes being less expensive are developed, the interest in these systems surely will increase.

It should be stated that using lead dioxide as anode material, overpotentials exceeding 1 V must be applied to generate oxygen. This implies a cell voltage exceeding 2.4 V, enough for the generation of ozone (see Sect. 5.4.2). The current efficiency concerning ozone of course is low (around 10%) and the specific energy use is high (80 Wh per gram of ozone) [55].

On the contrary, the production of ozone via gas discharge tubes needs only 14 Wh g^{-1}. Nevertheless, the electrolytic production line is paramount if only small amounts of ozone (<100 g h^{-1}) are

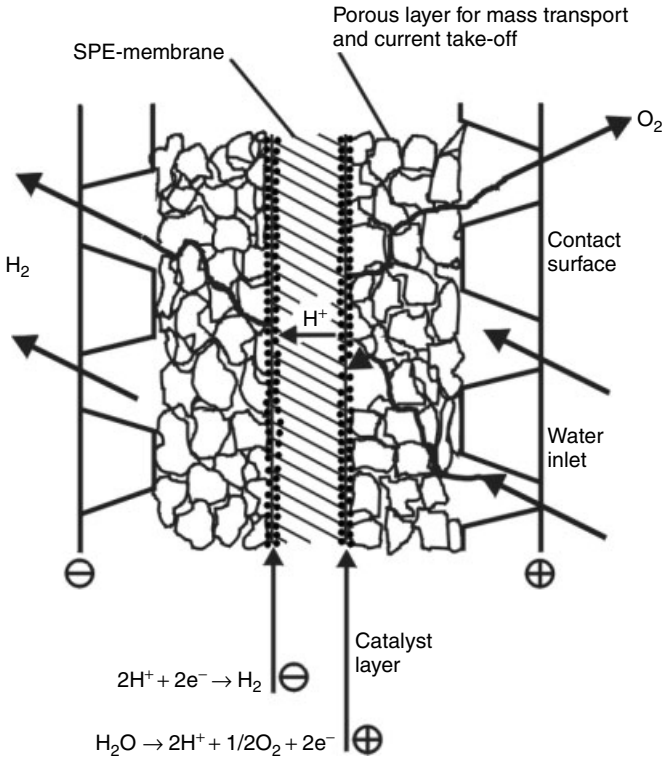


Fig. 23 Scheme for H_2O membrane electrolysis (not to scale). The contact elements are also the current collectors. The principle displayed originates from fuel cell technology, compare Chapter 8 of this volume.

needed (simple setup, no need for high voltage). Another advantage is the fact that ozone concentrations of 80 ppm in water or 200 g m^{-3} of gas are produced, values that can be achieved from gas discharge units only by further processing. The described electrochemical technique is commercially used in pure water technology [55].

5.4.3.3 Steam Electrolyzers

The temperature coefficient of the decomposition voltage of water is -0.85 mV K^{-1}, so at 100 °C this voltage amounts to 1.17 V. The coefficient for the decomposition of steam is -0.23 V K^{-1}. So, at 1000 °C the decomposition voltage would be 0.96 V. An electrolysis at 1000 °C can be carried out using ytterbium stabilized zirconium dioxide (YSZ, $ZrO_2 \cdot 0.15\ Y_2O_3$). This ceramic material shows O^{2-}-conductivity at temperatures exceeding ca. 700 °C ($\kappa = 0.5 \times 10^{-1} 1/\Omega cm$ at 1000 °C, compare Chapter 8).

The electrodes are directly attached to the surface of the ceramic material in the form of porous layers, consisting, for instance, of nickel (cathode material) or $LaMnO_3$ (an electron conducting mixed oxide as anode material). Current densities of 4 kA m^{-2} at cell voltages for water decomposition of only 1.3 V have become possible in this way [54]. Figure 24 shows the principle of a typical experimental setup. This (and corresponding flat configurations) is used in modern fuel cell technology, too (see Chapter 8).

Steam electrolyzers will have a long way to go before commercial applications are

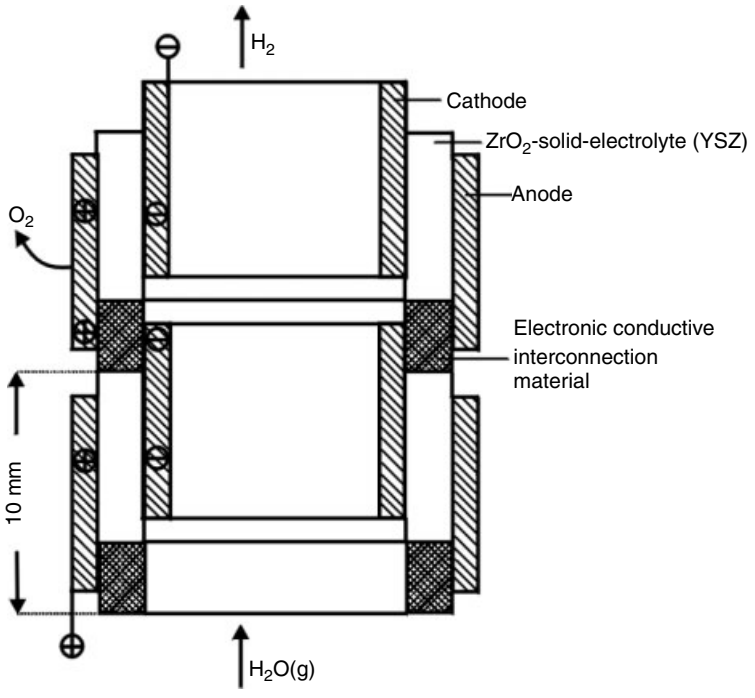


Fig. 24 Principle of steam electrolysis, compare text. Please note: the solid electrolyte has a thickness of approx. 0.5 mm.

in sight. This is due to the fact that problems concerning different coefficients of thermal expansion of the components and degradation of electrode coatings and interconnection materials are difficult to overcome at operating temperatures around 1000 °C. Possibly lowering of this value to 800 °C will be of help. But then the conductivity of the YSZ goes down, and so the thickness of the ceramic electrolyte has to be lowered to some 10 μm, a new problem.

5.4.3.4 Closing Remarks

Considering hydrogen economy, there are two aspects that have to be mentioned in addition.

First, in this scenario solar or wind power–operated water electrolyzers are included. These energy sources deliver their power discontinuously. So, the electrolyzers must be able to tolerate fluctuating loads, a goal to be achieved in present developments. Second, this economy includes the conversion of electricity to hydrogen by electrolysis and from hydrogen back to electricity via fuel cells. So, the efficiency of this sequence has to be taken into account when looking at the chances.

As a measure for this efficiency, clearly the cell voltage of the fuel cell compared to that of the electrolysis is a first tool. Nowadays we have values of 0.8–1.8 V, that is, the ratio is 0.45. In the future, 0.9–1.2 V may be achieved, that is, 0.75. But this value is somewhat optimistic, if the Faradaic losses (current efficiencies) and the energy losses resulting from electrolysis and fuel cell operation are taken into account, too, amounting surely to an additional loss of 10–15%.

So, if compared with the far less energy losses caused by electricity transportation, hydrogen economy is only an option. But even if it is not or only partially realized in the future, the advancement in water electrolysis technology is valuable.

Finally, it should be pointed out that the evolution of D_2 and HD is kinetically slower than that of H_2, leading to an enrichment of deuterium in the electrolyte. This has been an important preparative route for D_2 and HD.

5.5 Peroxodisulfates [7, 56, 57]

5.5.1 Introduction

Peroxodisulfates (of ammonium, sodium, or potassium; only these are of industrial importance) are used, for instance, in etching printed circuits, removal of photoresists and in organic polymer production, the anion being in general one of the most powerful oxidants. Worldwide, 80 000 tons of these species were consumed in 1998.

5.5.2 Electrochemical Reactions

Aqueous solutions of peroxodisulfate anions are manufactured by electrolysis only, that is, by anodic oxidation of sulfate or HSO_4^-

$$2SO_4^{2-} \longrightarrow S_2O_8^{2-} + 2e^- \qquad (36)$$

$$\text{or } 2HSO_4^- \longrightarrow S_2O_8^{2-} + 2H^+ + 2e^- \qquad (37)$$

($E_0 = 1.94$ resp. 2.06 V). The exact pathways are discussed controversially [58, 59]. A simple conception would be the intermediate generation of $SO_4^{\bullet -}$-radicals, which dimerize [24, p. 320]. The peroxodisulfate anions will form the salt with cations being present.

Considering the equilibrium potentials mentioned, the anode material must show a high overvoltage for oxygen evolution; considering the high acid and sulfate concentrations needed for the presence of the singly ionized species high chemical stability is necessary. Only platinum or platinum alloys (Pt/Ta) show the required properties. Nevertheless, the process is accompanied by the parasitic evolution of oxygen, even if overpotential increasing substances (like pseudo halides) are employed.

The cathode reaction is presented by hydrogen evolution from sulfuric acid. In this case, graphite may be used as electrode material.

5.5.3 Technical Processes

Figure 25 shows a single cell of a bipolar electrolyzer, consisting of 26 such elements, for the production of $(NH_4)_2S_2O_8$ from $(NH_4)_2SO_4$ as an example [60, 61]. Other cells used today origine from the former Weissenstein process (see next page). Their basic layout is described in [57].

The characteristic feature is a large number of vertical anolyte channels (9 × 3 mm), on whose back foils of platinum as active material are attached on tantalum foils mounted on a supporting plate. This is cooled from behind to realize an anolyte temperature of 30–45 °C (to diminish acid-catalyzed hydrolysis of the peroxodisulfate anion formed). Graphite acts as a cathode block, separated from the anode channels by a microporous polymer diaphragm, preventing reduction processes at the cathode.

Typical operational data are a cell voltage of 4–6 V at approx. 1 kA m^{-2}, a current efficiency of 80–90% and a specific energy consumption of 1000–2000 kWh per ton of $(NH_4)_2S_2O_8$. The $(NH_4)_2SO_4$ concentration in the anolyte changes from 380–440 g L^{-1} (inlet) to 200–240 g L^{-1} at the outlet, where the corresponding amount of $(NH_4)_2S_2O_8$ is found. 0.2 g L^{-1} $(NH_4)SCN$ is added as potential increasing medium. The catholyte consists of 300–700 g L^{-1} H_2SO_4.

It is possible to generate all three peroxodisulfates of industrial importance mentioned using direct electrolysis [62–64]. But in practice, the sodium

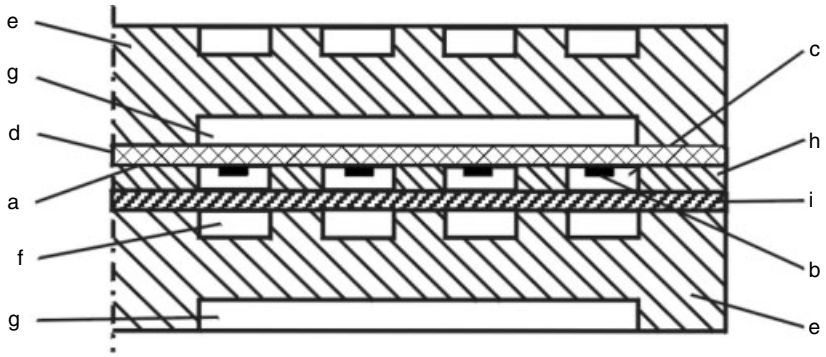


Fig. 25 Schematic view of a frame and plate electrolyzer for the production of peroxodisulfate. (a) Tantalum foil; (b) platinum foil strip anodes; (c) anolyte channel; (d) supporting plate; (e) graphite cathode block; (f) catholyte channel; (g) cooling channel; (h) spacer; (i) diaphragm.

and potassium salts are converted from the ammonium species. This is due to most favorable conditions with respect to solubilities of starting materials, products, and of current efficiencies in case of the ammonium species.

5.6 Hydrogen Peroxide [7, 57, 65]

5.6.1 Introduction

The main use of hydrogen peroxide is bleaching. Uses in chemical industry and in environmental protection are increasing because its great advantage is that the degradation product is water – a solution of some wt% is even used as a household cleaner.

In 1995, more than 2×10^6 tons (calculated as 100% H_2O_2) were produced.

5.6.2 Electrochemical Reactions

If reactions (36) and (37) are carried out without cations like NH_4^+ being present (compare Sect. 5.5.2), peroxodisulfuric acid will be produced from sulfuric acid. The produced acid can be hydrolyzed to give H_2O_2

$$H_2S_2O_8 + 2H_2O \longrightarrow 2H_2SO_4 + H_2O_2 \quad (38)$$

which can be separated by distillation.

This represents the classical (1908) electrochemical pathway for H_2O_2 – production (so-called Weissenstein process). Of course, for instance, $(NH_4)_2S_2O_8$ can be hydrolyzed, too, after being converted to potassium persulfate, (Münchner process) [66–68]. The drawbacks are high cell voltages (see preceding section) leading to an energy consumption of 13 000–17 000 kWh per ton of H_2O_2 (calculated as 100 wt%), as well as the necessity to use noble metals as anode material and a yield of only 70%. So, due to high costs this production line has become insignificant compared with catalytic processes developed later, such as the anthraquinone process.

But there is another possibility to generate peroxide by an electrochemical process, namely, the indirect reduction of oxygen (see this encyclopaedia Vol. 7, or [24, p. 281]). In both acidic and alkaline solutions, peroxide is formed as an intermediate. In today's practice, alkaline solution (NaOH) is used.

$$O_2 + H_2O + 2e^- \longrightarrow HO_2^- + OH^- \quad (E_0 = -0.076\ V) \quad (39)$$

$$HO_2^- + H_2O + 2e^- \longrightarrow 3OH^- \quad (E_0 = 0.88\ V) \quad (40)$$

The first two electron step can be made fast by using carbon material as electrode. This material on the other hand offers a high overvoltage for the second step and has a low catalytic activity for the decomposition of the peroxide ion. So, the net product is an alkaline solution of peroxide.

The necessity of a high overpotential for step (40) implies the use of low current densities. So electrodes with a high specific surfce are advantageous.

5.6.3 Technical Process

The principle can be carried out using gas diffusion electrodes [7]. Commercial reactors for the production of alkaline peroxide solutions have recently been realized using the principle of a monopolar trickle bed reactor [57, 69].

Figure 26 shows the principle used. The anolyte passes the diaphragm, which

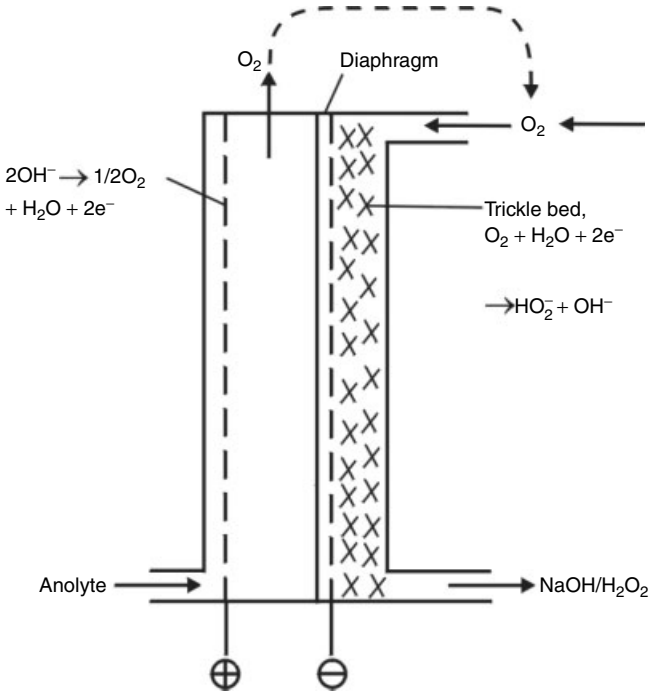


Fig. 26 Trickle bed cell.

serves as flow control element, and enters the trickle bed. This consists of graphite chips coated with a mixture of TeflonTM and active carbon, the chips being in the range of 0.5–2 mm in diameter.

In this way, hydrophobic regions are generated in the bed, so that it will not flood, ensuring an open path for the oxygen feed.

The commercial reactor contains single cells, realizing a production of 70 kg H_2O_2 per day, and they operate at a cell voltage of 2 V at a current density of 0.6 kA m^{-2}. The current efficiency being 90%, an energy consumption of 3.5 kWh kg^{-1} H_2O_2 is achieved.

Nevertheless, peroxide is least stable in alkaline solutions and so the production of pure H_2O_2 from the mixture is difficult and again uncompetitive considering the anthraquinone process. But the application of peroxide (see Sect. 5.6.1) actually requires alkaline solutions and concentrations that can be provided via Eq. (39) (approx. 5 wt%); that is, the cell liquor can be used directly in the bleaching process (on-site production).

Considering the cathode of a hydrogen–oxygen fuel cell, peroxide arises as an intermediate, too. Of course in case of a fuel cell, the accumulation of the intermediate is minimized by use of a suitable electrocatalyst. Changing this, the cell will produce peroxide [70]. In this case, no external power supply is needed.

5.7 Further Inorganic Species [71–73]

The electrochemical production of sodium peroxoborate tetrahydrate is no longer of

commercial importance and has not been included.

5.7.1 Chromic Acid Solutions

Aqueous chromic acid solutions are used, for instance, as pickling and chromium plating baths in metal and plastic processing. They can be produced by dissolving chromium(VI) oxide in water or from sodium dichromate(VI) and sulfuric acid.

Concerning the solution, the anions of the chromic acid (CrO_4^{2-}) and of dichromic acid ($Cr_2O_7^{2-}$) form an equilibrium

$$Cr_2O_7^{2-} + H_2O \rightleftarrows CrO_4^{2-} + 2H^+ \qquad (41)$$

In acidic solutions, $Cr_2O_7^{2-}$ (dichromate) is predominant, in alkaline media this holds for CrO_4^{2-} (chromate).

Chromic acid solutions can also be produced from acid chromium(III) solutions anodically (regeneration of chromium(III) solutions obtained from chromic acid solutions employed in oxidation processes).

$$2Cr^{3+} + 7H_2O \longrightarrow Cr_2O_7^{2-} + 14H^+ + 6e^- \qquad (42)$$

The standard redox potential of this process is 1.36 V. Equation (42) includes the possibility to regenerate Cr(III) solutions generated in using Cr(VI) as an oxidant. The cathodic reaction is

$$6H^+ + 6e^- \longrightarrow 3H_2 \qquad (43)$$

The electrolysis must be carried out in divided cells, otherwise freshly formed dichromate ions at the anode would be reduced again at the cathode. In industrial practice, a ceramic diaphragm is employed. The feed (acidic chromium(III) sulfate) is introduced first into the cathode compartment, where it becomes depleted of sulfuric acid (Eq. (43) and migration of sulfate ions into the anode compartment). Second, it is fed into the anode space, where the dichromate is formed and where the concentration of sulfuric acid is restored. Lost chromium is periodically replenished by adding chromium(III) oxide.

Owing to the extreme corrosiveness of the solution, the cell case must be of lead or must be lead lined. The anode must be of this material, too. This holds for the cathode, so high overvoltages for hydrogen evolution have to be taken into account.

So, the voltage of a single cell of a series arrangement is 3.5 V at 0.3 kA m^{-2} despite the fact that the anodic reaction is relatively unhindered. The current efficiency amounts to 80%.

5.7.2 Manganese Dioxide

Manganese is contained in many minerals; they generally contain more than 35% Mn, and consist of hydrated or anhydrous oxides. Only 5% of the world production of manganese ore is consumed in nonmetallurgy applications. From these 5% – amounting to 1.35×10^6 tons per year (1976) – 0.5×10^6 tons per year were associated with the manufacturing of dry cell batteries and a similar amount was used for the production of manganese chemicals. These data are probably still applicable, because worldwide ore consumption did not change remarkably.

Natural manganese dioxide used in dry cells must be of ''battery grade.'' This implies a content of 70–85% MnO_2 (44–55% Mn) in the ore, going along with a content of less than 0.05% concerning metals more electronegative than zinc, such as Cu, Ni, Co, or As (which

would corrode the zinc cathode of the dry cell). Other factors necessary for sufficient electrochemical activity are a suitable crystal structure, surface area, pore size distribution, particle size and shape, electrical conductivity, and so on.

All factors of course are of somewhat poor reproducibility concerning the natural MnO_2. So synthetic manganese dioxides are produced commercially for battery use (high performance cells): Chemical manganese dioxide (CMD) and electrolytic manganese dioxide (EMD).

EMD is the most important of the synthetic manganese dioxides (183 000 tons per year). The preparation starts from natural MnO_2, this being reduced to MnO firstly and then leached with sulfuric acid. The acidic manganese sulfate solution then is purified (precipitation of Fe, Pb, Ni, Co at pH of 4–6). After adjustment of concentrations to 75–160 g L^{-1} of $MnSO_4$ and 50–100 g L^{-1} of H_2SO_4, the solution is subjected to anodic oxidation

$$Mn^{2+} + 2H_2O \longrightarrow MnO_2 + 4H^+ + 2e^- \quad (44)$$

The cathode reaction is

$$2H^+ + 2e^- \longrightarrow H_2 \quad (15)$$

The electrolyzers usually are open steel troughs lined with rubber, ceramic, or plastic material. There exist different electrode configurations. A common configuration (used for instance in Japan, this country being the most important EMD manufacturer) follows the configuration of the electrowinning of zinc (see Chapter 4). In the present case, the anode plates are made of graphite, hard lead, or titanium. The latter being the modern trend, as Ti is mechanically and chemically stable and does not contaminate the product. But it must be mentioned that additives to the electrolyte are necessary to overcome possible passivation.

The cell voltage amounts to 2.2–3.0 V at 0.7–1.5 kA m^{-2}, the temperature is kept at 90–98 °C. To minimize evaporation, a paraffin layer on top of the electrolyte is common use. The current efficiency of modern plants can be higher than 90%.

If the MnO_2 layer deposited according to Eq. (44) at the anode surface reaches a thickness of 20–30 mm, it is removed mechanically and processed to a black powder showing the following features: typical content of MnO_2 (mostly the hexagonal ε–modification) 91 wt%; moisture 3–5 wt%; sulfate content up to 1.3 wt%; iron <0.02 wt%; Pb, Cu, Co <0.001 wt%; particle size <75 µm and BET surface 40–50 m^2 g^{-1}.

5.7.3 Permanganate

In the preceding section, the production of manganese chemicals was mentioned. These chemicals refer to the oxidation states up to VI. The preparation of potassium permanganate (oxidation state VII) is performed using electrolysis.

The procedure starts with MnO_2 ore finely ground, which is subjected to alkaline oxidative fusion, giving K_2MnO_4. From this species, the permanganate may be produced chemically by reaction with elemental chlorine or with carbon dioxide.

$$2K_2MnO_4 + Cl_2 \longrightarrow 2KMnO_4 + 2KCl \quad (45)$$

$$3K_2MnO_4 + 2CO_2 \longrightarrow 2KMnO_4 + MnO_2 + 2K_2CO_3 \quad (46)$$

But these processes have drawbacks: in the first case, half of the potassium is lost in form of KCl, in the second case one-third of the manganese fed in is found again in the form of MnO_2. In case of the electrolytic process, there are no such losses.

$$2MnO_4^{2-} \longrightarrow 2MnO_4^- + 2e^- \qquad (47)$$

Cathode reaction is

$$2H_2O + 2e^- \longrightarrow H_2 + 2OH^- \qquad (3)$$

using iron as electron material.

Electrolysis cells may be either monopolar or bipolar, rectangular, or circular and operated batchwise or continuously. Typical parameters are $KMnO_4^-$-content in the electrolyte 120–250 g L^{-1}, cell voltage 2.3–3.8 V at up to 1.5 kA m^{-2}, current efficiency up to 90%, operating temperature ca. 60 °C. The anodes consist of Ni or Monel. Figure 27 displays a circular electrolysis cell as an example.

The total volume of this cell is about 4 m^3, the electrode system is monopolar, and the operation mode is batchwise. The resulting $KMnO_4$ is allowed to crystallize within the cell and is drawn off periodically. As unfiltered electrolyte can be used in batchwise operated cells, the crude product must be recrystallized to free it from gangue.

As reported, potassium permanganate is made from MnO_2-containing ore by fusion followed by electrolysis. There exist, by the way, all-electrolytic methods. One

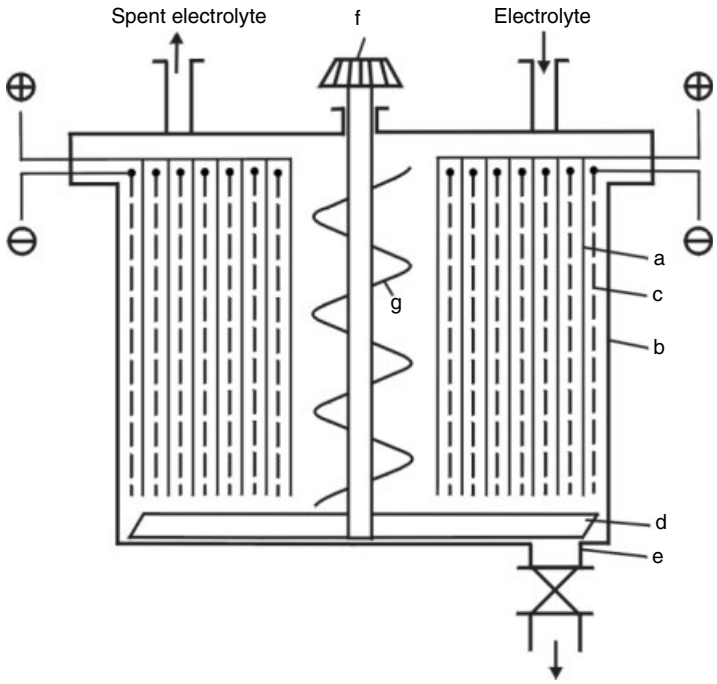


Fig. 27 Vertical section of circular permanganate cell (setup used in Bitterfeld, former GDR). (a) Anode; (b) rubber-lined trough; (c) cathode; (d) agitator; (e) discharge valve for product; (f) agitator shaft; (g) spiral agitator.

method involves direct anodic oxidation of ferromanganese [74], another possibility is the direct oxidation of MnO_2 in aqueous KOH [75].

Besides potassium permanganate, only sodium permanganate is of industrial significance. It can be made in several ways, including anodic oxidation of ferromanganese [76].

5.7.4 Copper Oxide

Copper(I) oxide, Cu_2O, can be used as seed and crop fungicide, as catalyst for numerous organic reactions (for instance hydrogenation of oils), or as starting material for the production of copper powders. Other examples are the use as antifouling pigment or as pigment for glasses or ceramic glazes.

The preparation of copper(I) oxide usually is carried out thermally or chemically using copper or copper compounds. But there exists an electrolytic pathway (of minor importance), too.

This pathway employs the anodic oxidation of copper metal to Cu^+ according to

$$2Cu \longrightarrow 2Cu^+ + 2e^- \qquad (48)$$

with an E_0 of 0.52 V, which preceeds the formation of Cu^{2+}. For electrode kinetics refer to this encyclopaedia, Vol. 7, and especially Reference 77.

The electrolysis is carried out in an undivided cell using copper electrodes and an aqueous solution of chlorides of alkali or alkaline earth metals. So copper (I) chloride, which is practically insoluble in pure water, will be formed and solved as a complex. At the cathode reaction (3) takes place

$$2H_2O + 2e^- \longrightarrow H_2 + 2OH^- \qquad (3)$$

The evolving hydrogen will provide for the transport of OH^--ions to the anode, so the Cu_2O is generated according to

$$2Cu^+ + 2OH^- \longrightarrow Cu_2O + H_2O \qquad (49)$$

This leads to the overall reaction

$$2Cu + H_2O \longrightarrow Cu_2O + H_2 \qquad (50)$$

The electrolysis employs a cell voltage of 1.2–1.4 V at 1.5 kA m^{-2} using a bath temperature of 70–80 °C. The bath concentrations amount, for instance, to 100–300 g NaCl/L and ca. 1.2 g NaOH/L. 0.8–1.3 kWh are needed to produce 1 kg Cu_2O. The resulting oxide appears as a suspension and must after filtration be stabilized by substances such as lime, gelatine or similar, due to easy oxidation by air.

References

1. P. Schmittinger, Th. Florkiewiez, L. C. Curlin et al., Chlorine in Ullmann's encyclopedia of industrial chemistry, 6th ed., *Electronic Release*, (Eds.: B. Elvers, S Hawkins), Wiley-VCH, Weinheim, 1999, pp. 185–279, Vol. 8.
2. Th. O'Brien, V. Bommaraju, F. Hine, *Handbook of Chlor-Alkali Technology*, Springer-Verlag, Heidelberg, 2004, Vol. 1–5.
3. P. Schmittinger, *Chlorine, Principles and Industrial Practice*, Wiley-VCH, Weinheim, 2000.
4. K. Blum, P. Schmittinger, Chlor, Alkalien und anorganische chlorverbindungen in *Winnacker-Küchler, Chemische Technik, Prozesse und Produkte*, Vol. 3, (Eds.: R. Dittmeyer, W. Keim, G. Kreysa, et al.), Wiley-VCH, Weinheim, 2005, pp. 430–544.
5. The Chlorine Institute Inc., *Chlorine Manual and various Pamphlets*, "L" Street, N.W. Suite 506, Washington D.C., p. 20036, 2001.
6. G. Paini, C. Tane, EuroChlor, Guidelines, Recommendations, Positions Papers, Ave. E. van Nieuwenhuyse 4, Box 2, B-1160 Brussels, eurochlor@cefic.be, www.eurochlor.org, 2003.

7. D. Hoormann, J. Jörrisson, H. Pütter, *Elektrochemische Verfahren – Neuentwichlungen und Tendenzen, Chemie Ingenieur Technik*, 2005, pp. 1363–1376 Vol. 77.
8. (a) Chemical Week, February 9, 2005, p. 28. (b) Chemical Week, February 19, 2003, p. 57.
9. H. Schultz, G. Bauer, E. Schachl et al., Potassium compounds in Ullmann's encyclopedia of industrial chemistry, 6th ed., *Electronic Release*, (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 281–328, Vol. 8.
10. H. Vogt, J. Balej, J. E. Bennett et al., Chlorine oxides and chlorine oxygen acids in Ullmann's encyclopedia of industrial chemistry, 6th ed., *Electronic Release*, (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 281–328, Vol. 8.
11. info@cellchem.ekachemicals.com, 2003.
12. S. Trasatti, G. Lodi, in *Elektrodes of Conductive Metal Oxides, part A*, (Ed.: S. Trasatti), Elsevier Scientific Publishing Comp., Amsterdam - Oxford - New York, 1980, p. 301 ff.
13. UhdeNora, Alkalichloridelektrolyse nach dem Quecksilberverfahren, Brochure.
14. C. H. Hamann, W. Vielstich, *Elektrochemie*, Wiley-VCH, Weinheim, 4. Auflage 2005.
15. T. A. Davis, J. D. Genders, D. Pletcher, Ion permeable membranes, Alresford, 1997.
16. UhdeNora, Company Information, Mai **2003**; Company Pamphlet, 2002; a) Asahi-Kasei, Company Information, July **2003**; Company Pamphlet, 2003; b) Chlorine Engineers Corp., Company Information, June **2003**; Company Pamphlet, 2003; c) www.eltechsystems.com/mcre.html, 2003; www.eltechsystems.com/3-1m2mcre. html, 2003; d) Electrochemical Technology Business, www.etbusiness.com/html/chlor-tech/fm_1500.htm; www.etbusiness.com/html/chlor-tech/bichlor.htm.
17. C. Kotzo et al., *Sulfate Removal from Brine*, presented at the Chlorine Institute Annual Meeting Houston, Texas, 2000.
18. G. A. Carlson, E. E. Estep, Porous cathode cell for metals removal from aqueous solutions, Paper presented at the ECS meeting, May 1972.
19. UhdeNora, Chlorine Symposium, 2004 see 4, p. 446.
20. EuroChlor Publication Environmental Protection, 11, Mercury Housekeeping, 4th ed., 1998.
21. B. S. Gilliatt, EuroChlor, Chlorine Industry Review, 2004–2005, 6; www.eurochlor.org/reviews.
22. L. C. Curlin, T. F. Florkiewicz, R. C. Matousek, Polyramix, A Depositable Replacement for Asbestos Diaphragms, *London International Chlorine Symposium*, London, 1988.
23. T. F. Florkiewicz, *The Future of Diaphragm Cell Technology*, Chlorine Institute Spring Meeting, St. Louis, Missouri, 1998.
24. C. H. Hamann, A. Hamnett, W. Vielstich, *Electrochemistry*, Wiley-VCH, Weinheim, 1998.
25. F. Gestermann, A. Ottaviani, *Modern Chlor-Alkali Technology*, Blackwell Science, 2001, p. 49 ff, Vol. 8.
26. G. Faita, F. Federico, UhdeNora Technologies SRL, WO 0157290, 2001.
27. F. Gestermann, Energiesparende elektrochemische Chlorherstellung mit Sauerstoffverzehrkathoden, Bayer Innovationsforum, 1999.
28. F. Gestermann, A. Bulan, H. Pinter, et al., DE 101 52 793 A1, 2003.
29. C. H. Hamann, P. Schmittinger, Degussa AG, EP 0 810 193 B1, 2002.
30. C. H. Hamann, J. Helling, P. Schmittinger, Degussa AG, EP 09 31 785, 2003.
31. C. H. Hamann, J. Helling, P. Schmittinger, Degussa AG, EP 10 18 499, 2003.
32. C. H. Hamann, T. Röpke, *J. Carbohydrate Chem.* **2005**, *24*, 13–17.
33. G. Huber, H. Pütter, K. Schierle, et al., BASF AG, DE 19859563, 2000.
34. F. Gestermann, *Modern Chlor-Alkali Technology*, Blackwell Science, 2001, p. 315 ff, Vol. 8.
35. C. Bienhüls, SGL Acotec GmbH, 27th International Exhibition- Congress, Abstracts of the Lecture Groups, Process, Apparatus and Plant Design, 2003, 252; C. Bienhüls, SGL Acotec GmbH, Company Communication, 2003.
36. www.kvaerner.com/chemetics/chemplant/sodchlor/process.htm, 2003; www.chemetics.ca, 2004.
37. Cellchem, Company Pamphlet, 2003.
38. P. Gallone, *Trattato di Ingegneria Elettrochimica*, Tamburini, Milano, 1973, pp. 487–498.
39. W. Müller, P. Jönck, *Chem. Ing. Tech.* **1963**, *35*, 78–80.
40. J. Jörrissen, *Ionenaustauscher- Membranen in der Elektrolyse und elektro-organischen Synthese., Fortschr.- Ber.*, VDI-Verlag, Düsseldorf, 1996 Vol. 3rd Series, No. 442.

41. C. H. Hamann, P. Schmittinger, J. Helling, Degussa AG (US), US 6294070, 2001.
42. M. Jaccaud, R. Faron, D. Devilliers, Fluorine in Ullmann's encyclopedia of industrial chemistry, 6th ed., *Electronic Release*, (Ed.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 379–393, Vol. 14.
43. A. Marhold, J. P. Joscheck, Anorganische Verbindungen des Fluors in *Winnacker-Küchler, Chemische Technik: Prozesse und Produkte*, Vol. 3, (Eds.: R. Dittmeyer, W. Kein, G. Kreysa, et al.), Wiley-VCH, Weinheim, 2005, pp. 637–642.
44. L. B. Asprey, U. S. Energy Research and Development Administration, US 3989808, 1976.
45. K. O. Christe, *Inorg. Chem.* **1986**, *25*, 3721–3722.
46. A. J. Rudge, in *Industrial Electrochemical Processes*, (Ed.: A. T. Kuhn), Elsevier, Amsterdam, 1971, pp. 1–69.
47. N. Watanabe, Y. Kanaya, *Denki Kagaku Kogyo Butsuri Kagaku* **1971**, *39*, 139–147.
48. P. Häussinger, R. Lohmüller, A. M. Watson, Hydrogen in Ullmann's encyclopedia of industrial chemistry, in *Electronic Release*, 6th ed., (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 85–240. Vol. 17.
49. J. Wolf, O. Weinmann, M. Bracha, Wasserstoff, in *Winnacker-Küchler, Chemische Technik, Prozesse und Produkte*, Vol. 4, (Eds.: R. Dittmeyer, W. Keim, G. Kreysa, et al.), Wiley-VCH, Weinheim, 2005, pp. 1109–1137
50. M. J. Kirschner, Oxygen in Ullmann's encyclopedia of industrial chemistry, in *Electronic Release*, 6th ed., (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 561–580. Vol. 24.
51. H. Wendt, *Electrochemical Hydrogen Technologies*, Elsevier Science Publishers, Amsterdam, New York, 1990.
52. V. M. Schmidt, *Elektrochemische Verfahrenstechnik: Grundlagen, Reaktionstechnik, Prozeßoptimierung*, Wiley-VCH, Weinheim, 2003.
53. M. Fischer, *Chem. Ing. Tech.* **1989**, *61*, 124–135.
54. G. Sandstede, *Dechema Monographien* **1992**, *125*, 329–355.
55. G. G. Scherer, H. Devantay, R. Oberlin, *Dechema Monographien* **1985**, *98*, 407–415.
56. B. Bertsch-Frank, D. Engel, P. Kleinschmit, Peroxo compounds, Inorganic in Ullmann's encyclopedia of industrial chemistry, 6th ed., *Electronic Release*, (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 433–462, Vol. 25.
57. G. Goor, E. Hägel, S. Jacobi et al., Peroxoverbindungen in *Winnacker-Küchler, Chemische Technik, Prozesse und Produkte*, Vol. 3, (Eds.: R. Dittmeyer, W. Keim, G. Kreysa, et al.), Wiley-VCH, Weinheim, 2005, pp. 675–758.
58. W. Smit, J. G. Hoogland, *Electrochim. Acta* **1971**, *16*, 821–831.
59. W. Smit, J. G. Hoogland, *Electrochim. Acta* **1971**, *16*, 961–979.
60. W. Thiele, DD 99548, 1972.
61. W. Thiele, K. Wildner, G. Heinze, Eilenburger Elektrolyse und Umwelttechnik GmbH, WO 9320261, 1993.
62. E. Rossberger, Peroxid Chemie GmbH, DE 2346945, 1975.
63. W. Thiele, K. Wildner, DD 119197, 1976.
64. A. M. O. Kärki, M. J. Lindström, H. T. Pajari, Kemira OY, DE 4426246, 1995.
65. G. Goor, J. Glenneberg, S. Jacobi, Hydrogen peroxide in Ullmann's encyclopedia of industrial chemistry, in *Electronic Release*, 6th ed., (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 253–284. Vol. 17.
66. W. Machu, *Das Wasserstoffperoxid und die Perverbindungen*, 2nd ed., Springer-Verlag, Wien, 1951.
67. W. C. Schumb, C. N. Satterfield, R. L. Wentworth, *Hydrogen Peroxide*, Reinhold Publ. Co, New York, 1955.
68. J. Balej, H. Vogt, *Fortschr. Verfahrenstech.* Sect. D **1984**, 22, pp. 361–389.
69. J. A. McIntyre, *Electrochem. Soc. Interface* **1995**, *4*, 29–33.
70. S. P. Webb, J. A. McIntyre, International Forum Electrolysis Chemical Industry 10th Electrosynthesis, Lancaster, N.Y., 1996, pp. 213–225.
71. J. Halstenberg, H. Knopf, P. Schmidt, Chromium compounds in Ullmann's encyclopedia of industrial chemistry, *Electronic Release*, 6th ed., (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 485–520. Vol. 8.
72. A. H. Reidies, *Manganese Compounds in Ullmann's Encyclopedia of Industrial Chemistry, Electronic Release*, 6th ed., (Eds.: B. Elvers, S. Hawkins), Wiley-VCH, Weinheim, 1999, pp. 543–564, Vol. 20.

73. W. Teworte, H.-J. Rabben, *Kupfer-Verbindungen in Ullmanns Enzyklopädie der technischen Chemie,* 4th ed., (Ed.: E. Bartholomé), Verlag Chemie, Weinheim, 1978, pp. 559–587, Vol. 15.
74. G. R. Popov, *Khim. Ind. (Sofia)* **1968**, *4*, 166–169.
75. T. Okabe, E. Narita, Y. Kobayashi, et al., Nippon Chemical Industrial Company Ltd., US 3986941, 1976.
76. H. Marcy, *Anorganische Manganverbindungen in Anorganisch-Technischen Verfahren,* (Eds.: F. Matthes, G. Wehner), VEB Deutscher Verlag, für Grundstoffindustrie, Leipzig, 1964.
77. M. Stelter, R. Holze, C. H. Hamann, *Bull. of Electrochem.* **1998**, *14*, 132–135.

6
Electrochemistry Applied to Organic Synthesis: Principles and Main Achievements

Jacques Simonet
University of Rennes, Rennes, France

6.1	**Introduction**	**321**
6.2	**General Feature of Electroorganic Synthesis**	**322**
6.2.1	Specificity and Prospects	322
6.2.2	Choice of a Cell	323
6.2.3	Choice of Electrode Material	323
6.2.3.1	Materials for Cathodes	323
6.2.3.2	Materials for Anodes	324
6.2.4	Main Reactions Achievable Electrochemically: A Survey of Experimental Conditions	324
6.3	**Some Specific Electrochemical Reactions Involving Carbon Atoms**	**325**
6.3.1	Building of Carbon−Carbon Bonds at the Anode and the Cathode	325
6.3.1.1	Kolbe Reaction: Anodic Oxidation of Carboxylates: Coupling of Aliphatic Free Radicals	325
6.3.1.2	Anodic Oxidation of Carbanions and Symmetrical Coupling	327
6.3.1.3	Anodic Oxidation of Aromatic Systems via Cation Radicals (Dimerizations, Trimerizations, Oligo, and Polymerizations)	327
6.3.1.4	Cathodic Coupling of Aryl Groups via Transition Metals Catalysis	330
6.3.1.5	Other Activation Reactions Involving Aromatic Halides in the Presence of Metal Catalysts	332
6.3.1.6	Aromatic Carboxylations	333
6.3.1.7	Electrosynthesis as a Tool to Afford Cyclizations	334
6.3.2	Anodic Substitution of Organic Derivatives Possessing a Mobile Hydrogen	335
6.3.2.1	Functionalizations of Nitrogen-containing Compounds	335
6.3.2.2	Oxidation of C−H Functions	337
6.3.2.3	Anodic Functionalization of Ethers	339
6.4	**Anodic Oxidation of Thioethers**	**339**

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

6.5	**Electrochemical Conversion of Main Organic Functions in Aqueous and Protic Organic Media**	**341**
6.5.1	Reduction of Unsaturated Compounds	341
6.5.1.1	Nitro Group	341
6.5.1.2	Reduction of Functions Possessing at Least One Nitrogen Atom	342
6.5.1.3	Reduction of the Carbonyl Group	344
6.5.1.4	Organic Acids and Derivatives	345
6.6	**Scission of Chemical Bonds by Means of Cathodic or Anodic Electron Transfers**	**346**
6.6.1	Main Cathodic Scission Reactions of Carbon−Heteroatom Bonds	346
6.6.2	Other Types of Bond Scissions	348
6.6.3	Application of the Electrochemical Scission of Bonds to the Chemical Deprotection	349
6.6.3.1	Anodic Deprotection of the Carbonyl Group: Cleavage of C−S Bonds	349
6.6.3.2	Amine Function Deprotection: Cathodic Cleavages of C−S and N−S Bonds	351
6.6.3.3	Protection and Deprotection of Alcohols and Thiols	353
6.7	**Electrosynthesis as a Tool for a New Solid-Phase Methodology**	**354**
6.8	**What is the State of the Art for Organic Silicon Derivatives?**	**356**
6.8.1	Electroreduction	356
6.8.2	Electrooxidation	358
6.9	**Some Words about the Use of Selenium in Organic Synthesis**	**358**
6.9.1	Electroreduction	358
6.9.2	Electrooxidation	359
6.10	**Electrochemical Activation of Small Inorganic Molecules Usable in Organic Synthesis**	**360**
6.10.1	Cathodic Intermediates from Dioxygen and Principal Types of Reactions	360
6.10.2	Reduction of Carbon Dioxide	361
6.10.3	Cathodic Behavior of Sulfur, Tellurium, and Selenium: an Easy Way of Forming Specifically Reactive Nucleophiles	361
6.10.4	Anodic Activation of Molecular Sulfur and of S−S Bonds	362
6.11	**Oxidation of Halides**	**363**
6.12	**Anodic Fluorination**	**364**
6.13	**Zero-current Electrolyses**	**365**

6.14 Elements of Electrochemical Chiral Synthesis 365

6.15 Some Models of Electrolysis Cells: From Laboratory Cells to Scaling up 366
6.15.1 Laboratory Cells 367
6.15.2 Industrial Cells 368

6.16 Possible Industrial Applications of Organic Electrosynthesis 369
References 371

6.1 Introduction

Electrochemical synthesis (organic, organometallic, or/and inorganic) principally rests on the capability of molecules, ions, and radicals to react at an electrode surface that facilitates electron transfer(s). In this regard, electrochemical synthesis is obviously based on the use of one reagent, the electron, abundantly (and cheaply?) furnished by the electricity supplier.

The electron (now considered as our main reagent) is, in principle, nonpolluting at least at the level of its use in electrosynthesis. In some countries where electricity is constantly produced by hydroelectric stations (hydroelectric power) or nuclear plants (atomic power), electricity can be inexpensive, especially at nights, weekends, and during vacations. However, any organic chemist who is convinced of the advantages of electrosynthesis must become familiar, before to setting up an electrolysis cell, with the sign of electrode, the direction of the current flow, as well as with the very basic units like the ampere, the volt, and the coulomb. Anyhow, everyone has heard about analytic electrochemical methods such as voltammetry and coulometry. In particular, a voltammetric scan can be regarded as a microreduction (or oxidation) within the diffusion layer around the working microelectrode (made principally of platinum, mercury, or glassy carbon). The interest of the analytical method is to obtain preliminary but essential information. This information includes the information concerning macroelectrolysis, which is the desirable mode of electrosynthesis, the potential to be applied, the influence of the solvent (pH, intrinsic properties of the medium), and the nature (material) of the working electrode. Other parameters that need to be examined include the cell geometry, temperature effects, substrate concentration, and so on.

In order to take into account carefully the analytical data concerning the electrochemical reaction to be carried out, two basic equations must be recalled:

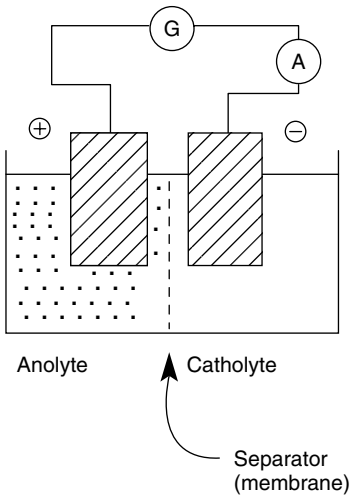
- The potential difference at the cell terminals given by

$$\Delta E = |E_{i=o}| + \eta_a + |\eta_c| + IR$$

were $E_{i=o}$ accounts for the potential difference at zero current, which can be calculated by application of the Nernst equation when the redox couples taking place at each electrode are clearly identified, η_a and η_c are the overvoltages, respectively, at the anode and the cathode, and IR is the ohmic drop due to the electrical resistance of the electrolytes and the

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

separator;



- The Nernst equation takes into account the thermodynamics of the two redox reactions (fast in the sense of the kinetics of the heterogeneous electron transfer, but here in the absence of associated chemical reactions). The thermodynamic properties of a redox reaction are described by the standard potentials, E°, and the concentrations of the redox species at the electrode surface. At the anode, we have an electron extraction from the electroactive species, and at the cathode the electron is donated to an electroactive species to effect a reduction. These processes may be represented in the general form as:

Oxidation

$$\mathrm{A} \underset{+ne^-}{\overset{-ne^-}{\rightleftharpoons}} \mathrm{B} \qquad E = E^\circ_{\mathrm{AB}} + \frac{RT}{nF} \ln \frac{[\mathrm{B}]_s}{[\mathrm{A}]_s}$$

Reduction

$$\mathrm{C} \underset{-n'e^-}{\overset{+n'e^-}{\rightleftharpoons}} \mathrm{D} \qquad E = E'^\circ_{\mathrm{CD}} + \frac{RT}{n'F} \ln \frac{[\mathrm{C}]_s}{[\mathrm{D}]_s}$$

6.2 General Feature of Electroorganic Synthesis

6.2.1 Specificity and Prospects

Electrochemistry should appear as a very specific field in which activation is essentially based on electron-transfer reactions at a level of energy (here, the applied electrochemical potential of the concerned electrode) that is fully controlled. Of course, such electron transfers occur heterogeneously at electrodes (generally metals or solid electronic conductors, such as various types of carbon) that can be chemically modified in order to specifically direct chemical reactions (occurring before or after the first charge transfer). Generally, the activation of molecules (or ions) via electron transfer often corresponds to significant energies. From organic substrates mostly entities possessing an odd number of electrons are formed. Depending on the structure of the starting compounds (molecules or organic ions), essentially radical anions, radical cations, or radicals may be produced. The existence of such transient species can be followed by the electron spin resonance technique, at least when the reactions associated with (and/or subsequent to) the electron transfer are not too fast.

When aqueous solutions (buffered or unbuffered) are used (and their pH adjusted to a desired process), electrolytic reactions afford hydrogenations at cathodes and oxidations of the diverse functional groups at anodes. In other words, cathodic saturation of double bonds or carbonyl groups is, in principle, achievable under these conditions, while alcohols are anodically transformed into ketones or acids. Additionally, catalytic hydrogenations can also be performed in acidic electrolytes at

specific metallic electrodes like platinum or palladium.

6.2.2 Choice of a Cell

The design of the cell determines the success of the electrochemical process. Very generally, most of cells possess a separator (porous glass or organic membrane) between the anodic and cathodic compartments. For achieving experiments on the laboratory scale (0.1 to 10 g), even rudimentary geometries can be quite efficient. It must also be emphasized that a cell giving the best results for one reaction might not be the one well adapted for another.

For laboratory use, most electrolyses can be achieved with a U-shaped cell equipped with a glass frit separator. In electrochemistry, the notion of the working electrode dipping in a well-defined surrounding is essential. The electrochemist defines the reaction (oxidation at the anode in the anodic compartment or reduction at the cathode in the cathodic compartment) to be achieved, and what happens at the other electrode (called *counter electrode*) does not really matter (although one has to avoid efficiently any diffusion of effluents formed in the other compartment). Some examples of electrolysis cells are shown and commented on at the end of this chapter (Sect. 6.15).

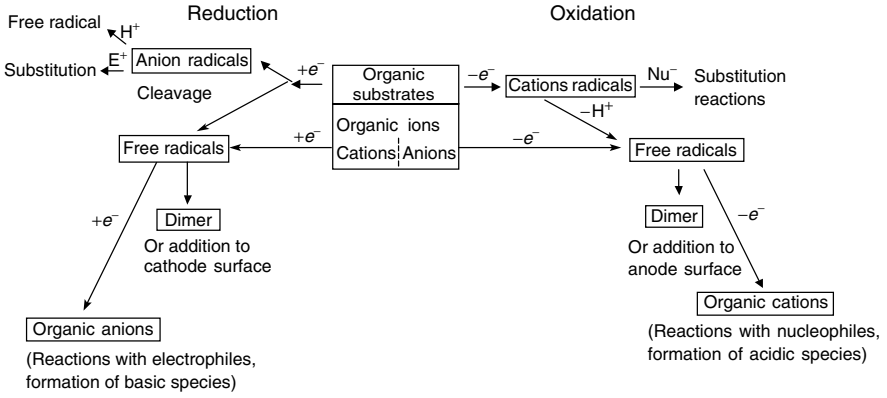
On the laboratory scale, it may be helpful to use three-electrode cells, where the electrochemical potential of the working electrode is set against that of a reference electrode, itself connected to a potentiostat. The purpose of a potentiostat is to fix precisely the reduction (or oxidation) potential necessary for the complete conversion of an organic substrate. At the end of the reaction, the electrochemical current has totally vanished and the concomitant use of a current integrator (or coulometer) may allow one to apply Faraday's law to organic compounds (knowledge of the number of electrons implied in the considered oxido-reduction process).

In the case of well-known electrochemical reactions, as well as for electrolyses where larger scales are involved, two-electrode cells (connected to a galvanostat) can be used with continuous feed of the reactant to the working electrode. This type of electrolysis is suitable for industrial purposes where specific devices and cells are utilized. Since electrodes of large areas are necessary, the distance between the anode and the cathode is small and determines the cell geometry (e.g. capillary-gap cell or filter-press cell). The use of cells equipped of porous electrodes (materials like graphite or carbon moss, platinum, nickel) have also been developed to perform electrocatalytic reactions at very large surfaces. Some typical cells used in the laboratory and in industry are presented at the end of this review.

6.2.3 Choice of Electrode Material

6.2.3.1 Materials for Cathodes

The use of mercury is – in principle – forbidden on the bench for environmental reasons. It is a pity, because mercury permits one to reach very high reducing potentials (−3 V versus the saturated calomel electrode (SCE), which is about what is necessary to achieve the reduction of benzene under conditions near to the Birch reaction). Substitute materials can be glassy carbon and graphite. Platinum and gold plates or mosses are of interest but their use is limited (because of the occurrence of hydrogen evolution) in acidic or aqueous solvents.



Scheme 1 Principal reactions (via electron transfer) in organic electrochemistry.

6.2.3.2 Materials for Anodes

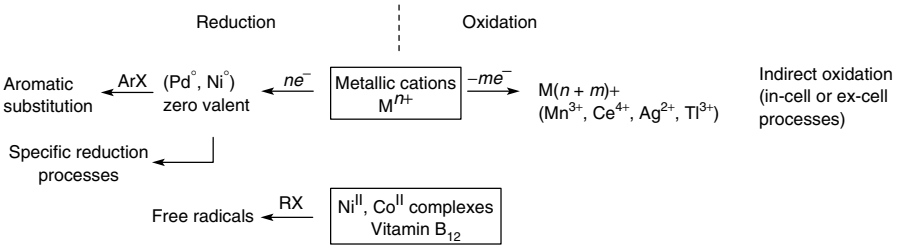
Platinum and glassy carbon are the quasi-universal materials for most anodic oxidations, irrespective of the solvent and conditions required to achieve electrolysis. Carbon and graphite felts can be employed as well. Most of these materials are commercially available from Japanese and American firms involved in the production of carbons.

6.2.4 Main Reactions Achievable Electrochemically: A Survey of Experimental Conditions

It is noteworthy that the use of organic solvents allows one to considerably widen the synthetic possibilities of the electrochemical technique. In particular, electrolyses performed in aprotic polar solvents (e.g. dimethylformamide, acetonitrile, methylene chloride, dimethylsulfoxide, propylenecarbonate) can enhance synthetic possibilities to provide a convenient way of forming, upon electron-transfer activation, of (1) nucleophiles (radical anions, anions), (2) electrophiles (radical cations, cations), and (3) free radicals (e.g. by means of oxidation of anions or scission of radical anions) when these solvents can be considered as poor H donors.

Cleavage reactions of radical ions may be conveniently applied in deprotection processes like the cathodic scission of sulfonamides (deprotection of amines) and the anodic cleavage of gemdithioacetals (deprotection of carbonyl groups).

The utilization of poorly conducting solvents is possible in electrochemistry,



Scheme 2 Use of inorganic cations in some indirect cathodic and anodic processes.

but only in the case where the electron transfer occurs to generate free radicals (induction of radical chains, e.g. using Ni^{II} or Co^{II} complex catalysts) or charged radicals (reactions of cycloadditions or cyclodimerization under the conditions chosen, so as to keep these reactions catalytic in electron). Reactions involving zero-valent metals can be useful for processes leading to the formation of carbon–carbon bonds from organic halides.

The ease of formation of free radicals brings chemists to consider electrochemistry as a convenient method to afford dimers. Obviously, reactions of this type are totally stoichiometric and demand 1 mol of electron per mole of a substrate.

In some cases, especially in oxidation, the use of mediators by the electrogeneration of high-valency metallic cations can be helpful. On the other hand, if similar, indirect reactions are possible in the reduction range, organic mediators have to be used. Such redox catalysis reactions are perturbed by the fact that couplings may occur between π-acceptor anion radicals used as electrogenerated reducing species and a free radical issuing from the cleavage reactions. Schemes 1 and 2 summarizes the principal reactions achievable by electrochemistry. Obviously, the key words are radicals, charged radicals, electrogenerated acids (EGA) and bases (EGB), nucleophiles, and electrophiles, catalysis by zero-valency metals or by valency change of metallic cores of metallo complexes.

The use of specific, electrochemically charged interfaces also appears to be of utmost importance, since structure and properties of these conducting materials can be changed with an applied potential.

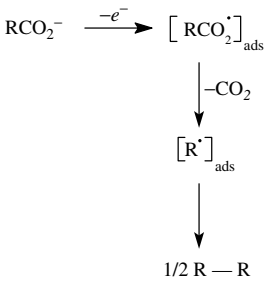
In summary, the electrochemical technique is in fact quite simple but is also rich in possibilities. Recent exhaustive reviews and books [1] could guide the reader in the case of very specific problems. The following pages are aimed at drawing our minds to what electrochemistry is really able to achieve, even though many specific reactions conducted at the electrodes are not mentioned here.

6.3 Some Specific Electrochemical Reactions Involving Carbon Atoms

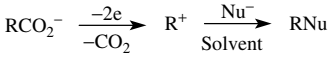
6.3.1 Building of Carbon–Carbon Bonds at the Anode and the Cathode

6.3.1.1 Kolbe Reaction: Anodic Oxidation of Carboxylates: Coupling of Aliphatic Free Radicals

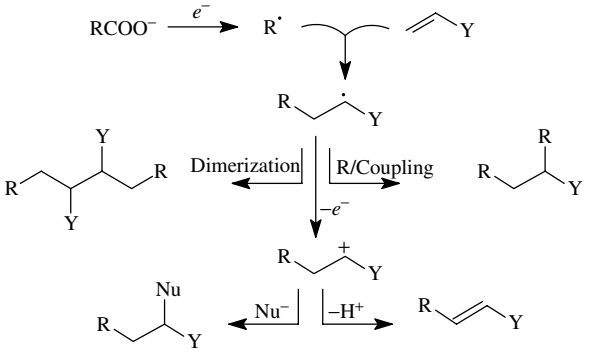
The anodic oxidation of carboxylates at a platinum electrode (high concentration, high current density, mild experimental conditions) leads to the formation of adsorbed radicals (via a decarboxylation process), which dimerize to a corresponding hydrocarbon:



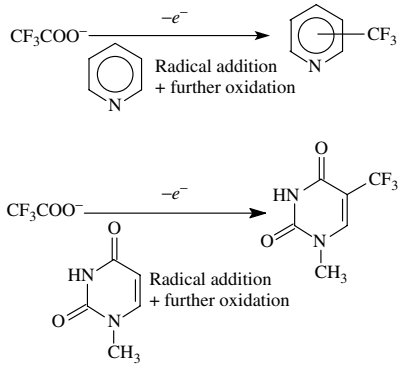
The use of other electrode materials (like graphite) allows the oxidation of the transient radical as a ''non-Kolbe'' reaction. The reaction turns out to be a two-electron process. The nature of products depends then on the nucleophilicity of the solvent.



The Kolbe reaction can be successfully used [2, 3] for the generation of free radicals which dimerize or add onto unsaturated systems that are essentially radicalophile. The product distribution will depend on the nucleophilicity of the solvent/electrolyte, on the pH, as well as on the kinetics of different competing pathways in the following scheme.



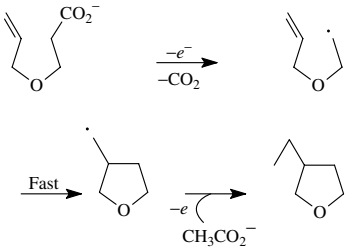
The ease of radical generation from carboxylates was exploited in view of the preparation of $CF_3^{\bullet}$ radicals from the oxidation of trifluoroacetates (it is worth noting that $CF_3^{\bullet}$ can also be obtained from a one-electron reduction of CF_3X, X = halide such as Br, Cl, or the arylsulfonyl group).



However, many examples in the literature concern mixed electrolyses of carboxylates that can afford dissymmetric dimers (in mixtures with symmetric ones)

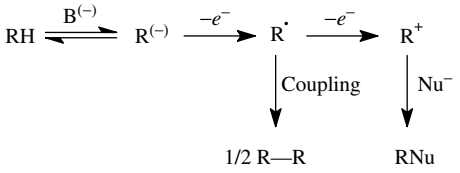
$$R'COO^- + RCOO^- \xrightarrow{-e} R'-R$$

Reactions of this type are interesting when any cheap and trivial carboxylates are considered. Thus, the electrolysis of high carboxylic acids in the presence of a large excess of acetate ions provided [4] methylated products in high yields. Moreover, examples of radical addition corresponding to internal rearrangements may exist. Thus, in the scheme below, the oxidation of the transient radical (after internal cyclization) was not observed in the presence of a large amount of methyl radicals concomitantly produced at the anode surface (electrolyte: sodium acetate).



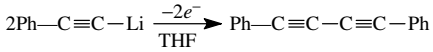
6.3.1.2 Anodic Oxidation of Carbanions and Symmetrical Coupling

Carbanions (those with the largest use are Grignard reagents, organolithians, or anions at the vicinal position to an unsaturated system) are easily oxidized. The anode material may be platinum, vitreous carbon, graphite, as well as titanium oxide.

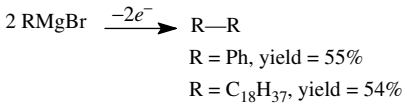


The yield of the dimer mainly depends on the experimental conditions (concentration, current density, electrochemical stability of R• at the potential of its formation (E°_{RH} should be smaller than $E^{\circ}_{R^+/R\bullet}$)). Many examples of dimerization [1] can be drawn from the recent literature:

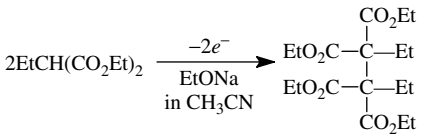
- dimerization of stabilized anions from arylacetylenes



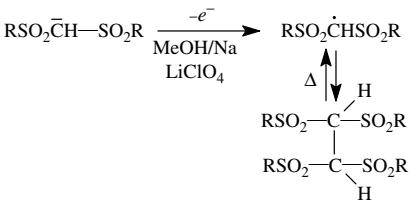
- anodic coupling of Grignard reagents



- dimerization of malonates



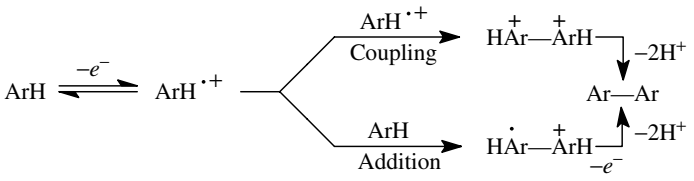
- coupling of the anions of disulfones [5, 6] (allowing a reversible radical dimerization at the temperature applicable for the activation of radical chain processes)



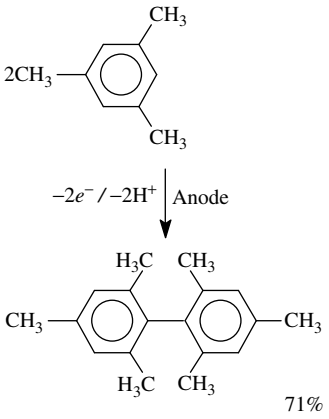
(if R = Et, yield: 78% at room temperature)

6.3.1.3 Anodic Oxidation of Aromatic Systems via Cation Radicals (Dimerizations, Trimerizations, Oligo, and Polymerizations)

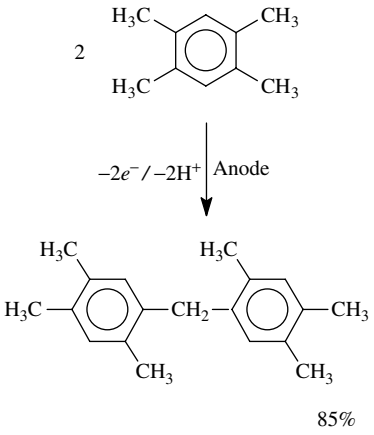
Generally, in non-nucleophilic solvents (yet in the presence of non-nucleophilic organic bases like lutidine or inorganic bases like sodium carbonate) two ways of couplings may concern electrogenerated aromatic cation radicals. Thus, with an aromatic substrate such as ArH, radical coupling [7] may compete with the addition cation radical/substrate.



It must be kept in mind that such symmetrical couplings are not inevitably favored. Thus, mesitylene yields (platinum anode, acetonitrile as a solvent containing tetrabutylammonium tetrafluoroborate) an aryl–aryl coupling product,

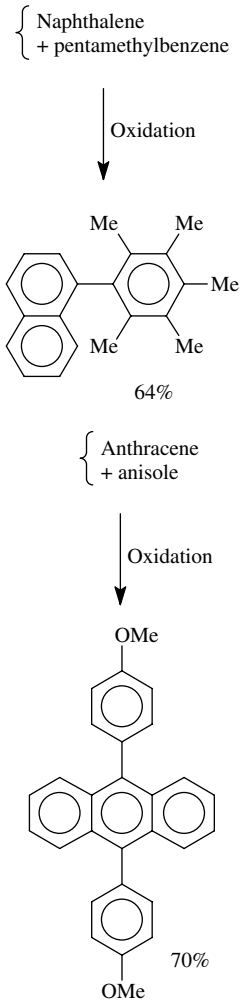


while 1,2,4,5-tetramethylbenzene [8] essentially affords (carbon anode, methylene chloride as a solvent, and nBu$_4$NBF$_4$ as a supporting salt) a coupling that involves a methyl group.



The coupling regioselectivity appears to be strongly dependent on the acidity of the transient cation radical and the activity of the intermediary cation in electrophilic substitution reactions. However, in the two examples quoted above, the hindrance argument cannot explain totally satisfactorily the change in chemoselectivity.

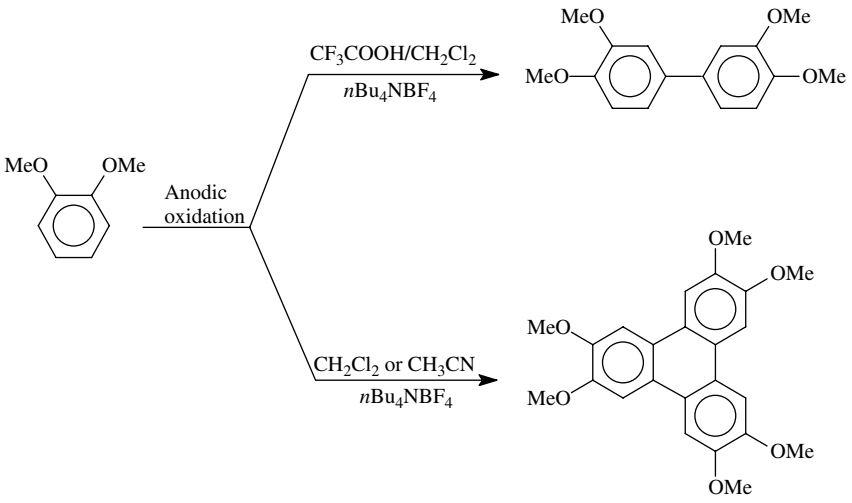
As a matter of fact, mixed anodic couplings with substituted aromatic structures have been shown to be possible:



With aromatic systems substituted by π-donor groups (e.g. amino ether and thioether groups), coupling reactions were found to be more complex because dimers generated at the anode are – often – more

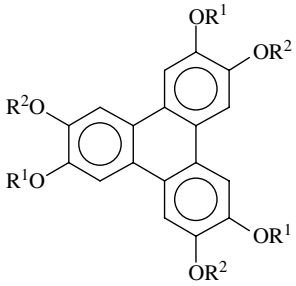
electroactive than the monomer itself ($E^{o}_{Monomer} > E^{o}_{Dimer}$). This results in a loss of selectivity for dimerization, since the cation radical of the dimer may independently lead to side reactions. The ease of oxidation of dimers is due to an increase (through better conjugation) of their HOMO (highest occupied molecular orbital) energy levels. In order to obtain dimers in good yields, two strategies are possible: (1) the use of flow cells equipped with a porous anode from which the oxidized form is instantaneously removed from the oxidation zone and then reduced chemically or electrochemically at the bottom of the device (e.g. electrogenerated superoxide or zinc powder); and (2) the obtaining, under well-chosen experimental conditions, of a mixed-valency salt from the dimer radical cation as a conducting deposit on the solid anode. This deposit can further be reduced according to conventional chemical (and electrochemical) procedures.

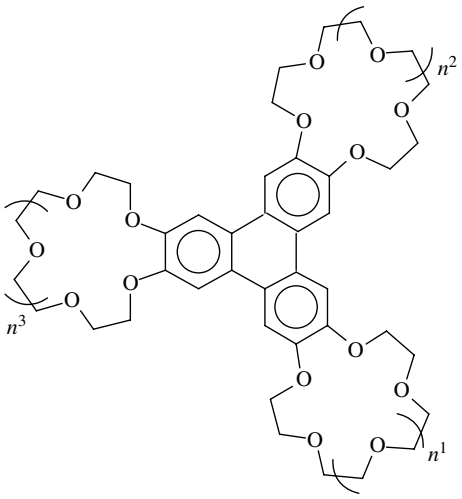
An example worth mentioning is the synthesis of trimers possessing polyether chains or crown ethers. More precisely, veratrole taken as a model molecule may dimerize or trimerize according to experimental conditions.



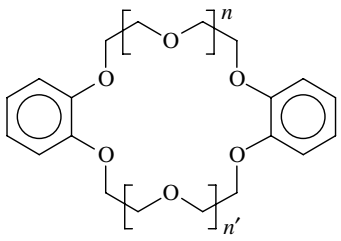
Let us note that the dimer (see scheme above) is only an intermediate for the formation of the trimer. However the relative ratio of these two products strongly depends on the experimental conditions. More precisely, conditions that favor trimerization [9] concomitantly with the electrocrystallization at the anode (mainly platinum) were successfully found and then applied for electrogenerating liquid crystals possessing discotic or/and ionophoric properties. We recall that a conventional organic synthesis does not allow easily such molecular engineering, since the oxidizing species are mainly high-valency salts (e.g. $FeCl_3$), which also play the role of Lewis acids. Accordingly, the electrolysis of mixtures of ortho diethers (R^1 and R^2 = long alkyl chains or polyether chains) easily affords discotic systems that are endlessly [10] modulated. Additionally, triphenylenes with an ionophoric

behavior (possessing 1, 2, or 3 crown ether moieties) may be advantageously synthesized.





Lastly, such a trimerization reaction – until now specific for orthodialkoxybenzenes – could be applied [11] to the polymerization (as polytriphenylenes) of dibenzocrown ethers leading to resins supported by conducting materials. The ionophoric properties of such products have been used for the extraction of heavy metal cations from water when the concentrations of these cations are extremely small.

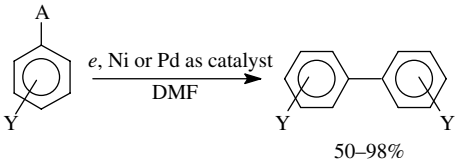


6.3.1.4 Cathodic Coupling of Aryl Groups via Transition Metals Catalysis

Practically, electrochemistry does not show any obvious examples of coupling of electrogenerated aryl radicals. Such a coupling involving antibonding orbitals is not favored and, in addition, radicals of this kind are produced, often at rather negative potentials, that render possible their reduction into anions.

The group of Nedelec in France has demonstrated that the use of transition metals (electrogenerated Ni°, Pd°, Zn°, etc.) may elegantly afford the coupling of a very large palette of aromatic compounds resulting in a carbon–carbon bond formation.

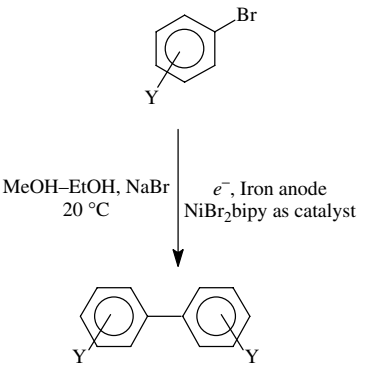
In the field of symmetrical aromatic coupling, this method is particularly interesting with halides and triflates [12–14].

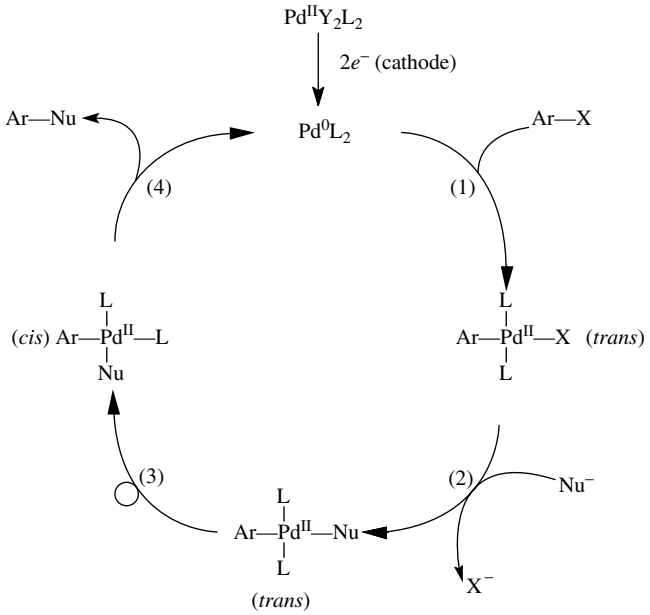


A = I, Br, Cl, CF_3O_2

Y = alkyl, aryl, MeO, CN, CF_3, NMe_2, MeCO, etc.

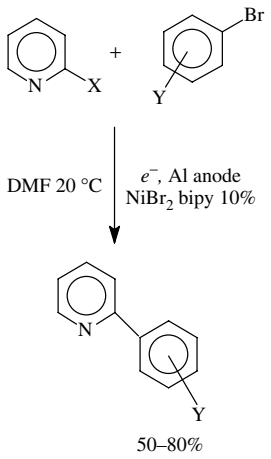
The mechanism was discussed recently [15]. The zero-valent metal (produced at the interface by reduction of the metal cation) induces an oxidant addition, in principle followed by nucleophilic substitution and a reducing elimination. The scheme below exhibits a simplified catalytic mechanism with aromatic halides. Electrochemistry is involved in the activation phase (formation of zero-valent metal [16–18]). The nucleophile is obviously produced by the oxidant insertion, and in the catalytic cycle hereafter Ar—Nu has to be considered of course as the Ar—Ar dimer. most of cases, ligand L is ortho, *ortho*-piperidine(bipy)).

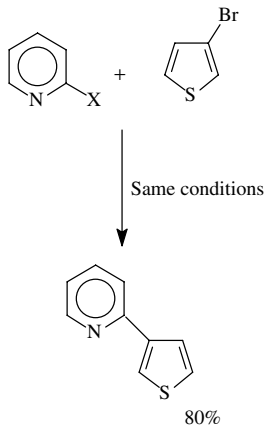




Other examples can be given concerning the coupling of arylbromides [19] with the use of nickel metal catalysts. Reactions of this type may be achieved in protic media (20:80 Dimethylformamide (DMF)/ethanol or 50:50 ethanol/methanol mixture) with yields comparable with those usually found in aprotic media. (In

Formation of aryl–aryl links was demonstrated to occur between aromatic bromides and 2-chloro or 2-bromopyridines. These reactions are efficient, since they allow in one step substitution reactions that are obviously much more difficult to achieve by the usual chemical methods.

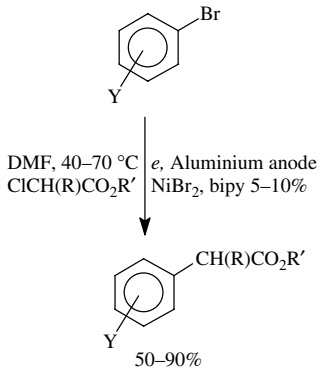




6.3.1.5 Other Activation Reactions Involving Aromatic Halides in the Presence of Metal Catalysts

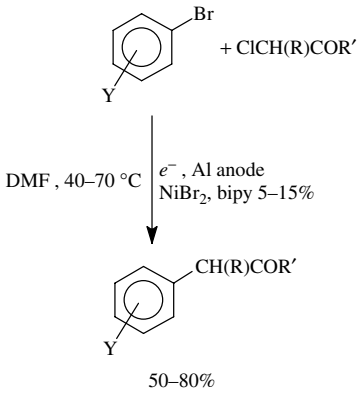
Reactions of heterocoupling have also been widely studied and the synthetic possibilities appear to be of high interest. (In the scheme below, A stands for an activating group.)

The route (a) was explored [20] with α-halogenoesters ($ClCH_2CO_2R$ or $ClCH(CH_3)CO_2R$): it yields α-arylacetic or α-arylpropionic esters.

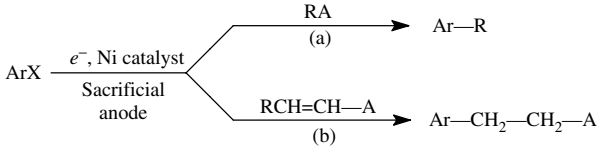


(with Y = CH_3, CN, F, CF_3, Me_2N, COMe, etc)

With α-chloroketones, benzylketones can be conveniently formed [21]:

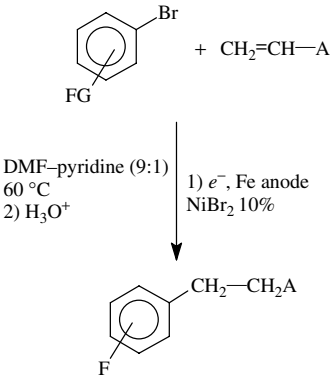


In the two latter cases, the unavoidable diaryl compound formation is minimized by the progressive and slow addition of the ''activated'' RCl compound to



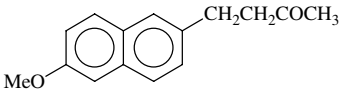
the ArBr/DMF solution. Electroactive Y groups are not converted by such electrocatalytic reactions. Substitution reactions were shown to be also effective with 2- and 3-bromothiophene (use of α-chloroesters and α-chloroketones).

On the other hand, creation of carbon–carbon bonds from ArX by addition to activated olefins could be achieved [22].

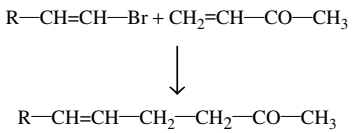


with FG = CN, OCH_3, $COCH_3$ and A = CO_2Et, CN, $COCH_3$.

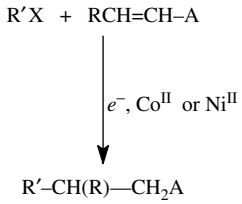
This reaction was particularly applied [23] to the one-step synthesis of nabumetone



with an isolated yield of 90% based on the starting bromonaphthalene. The method was also applied to the addition to vinyl groups [24] with a good stereospecificity. The use of acetonitrile as a cosolvent was preferred.



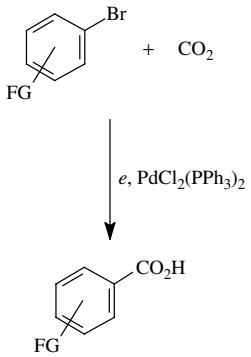
Lastly, alkylation catalyzed by Co^{II} complexes (like vitamin B_{12} or cobaltoximes) or by Ni^{II} cyclam can be efficiently completed [24–26]. In such cases (inner sphere), the electron transfer occurs via electrogenerated Co^{I} or Ni^{I} complexes.



It is important to stress that the electrochemical reaction is here fully stoichiometric. The term ''catalytic'' used for Ni or Co intermediates (constantly regenerated in the course of the electrochemical reaction) has therefore to be carefully differentiated from reactions using Ni^{II} or Pd^{II} as listed above.

6.3.1.6 Aromatic Carboxylations

The electroreduction of aromatic halides in the presence of catalytic amounts of [Ni–PPh_3] [27] or [Ni–dppn] [28] mainly yields arylcarboxylates. Such a reduction process may also be catalyzed by palladium complexes.



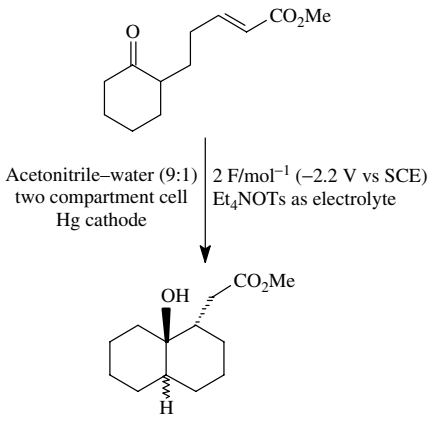
FG = Cl, CF_3, CO_2Et, F, CH_3
(FG) functional group

6.3.1.7 Electrosynthesis as a Tool to Afford Cyclizations

Intramolecular reactions can be considered as a powerful tool in organic synthesis to permit the building of carbon–carbon bonds in a regio- and stereoselective manner [29]. Thus, by means of electrochemistry several ways can be envisaged [30] to create carbon–carbon bonds, essentially by: (1) intramolecular coupling of radical ions (cations or anions at the anode and the cathode respectively); (2) the formation of free radicals (by cathode bond cleavage); (3) the use of Ni^{II} or Co^{II} catalysts followed by addition onto unsaturated systems belonging to the molecule (here, the triggering of the reaction – often a radical chain – occurs via the generation of the free radical at the cathode; (4) formation of a nucleophile (mainly by a two-electron reduction) capable of reacting in an intramolecular way by attacking any electrophilic site; and (5) Kolbe generation of a transient σ radical at the anode.

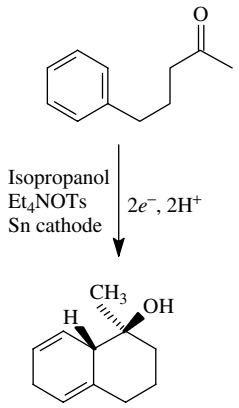
A few examples of monomolecular cyclizations are given below:

- via formation [31] of a nucleophile located on the more electroactive group;



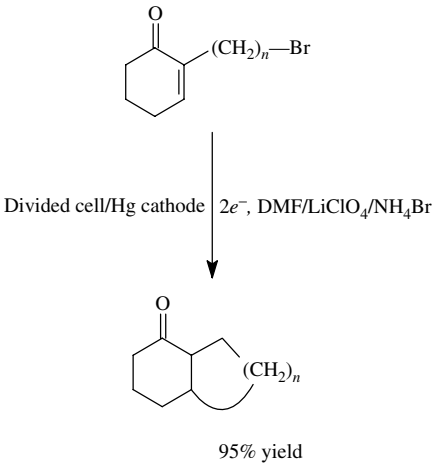
trans/*cis*: 2.5/1

- cathodic activation [32] presumably of the aromatic moiety at very negative potentials with an attack of the electrogenerated nucleophile on the carbonyl group;



cis/*trans*: 100 : 1

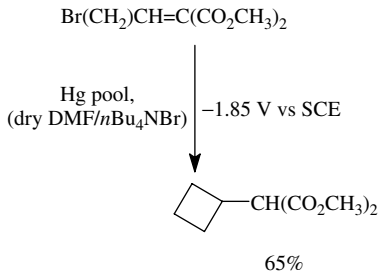
- use of aquocobalamine (intermediates: alkyl-Co^{III} complexes) in the cyclization [33] of substrates possessing ω-bromoalkyl functions;



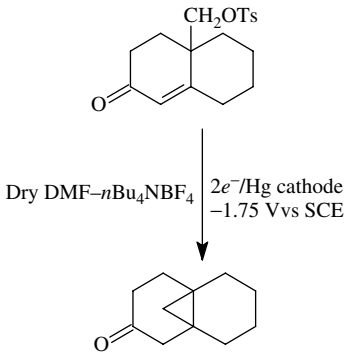
95% yield

- electrogenerated nucleophile formation from the π-acceptor through a two-electron reduction followed by a bromide elimination. This is the case [34]

for dimethylcyclobutylmalonate;

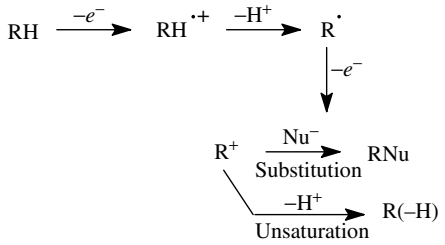


- formation [35] of a cyclopropane ring similar to a nucleophile formation/elimination process.



6.3.2 Anodic Substitution of Organic Derivatives Possessing a Mobile Hydrogen

In a general way, although many exceptions can exist that depend on intrinsic structures of organic substrates, the anodic oxidation of functional groups with mobile hydrogen is depicted as follows (the process implies a radical cation as primary intermediate, whose acidity is strongly enhanced compared to the starting molecule):

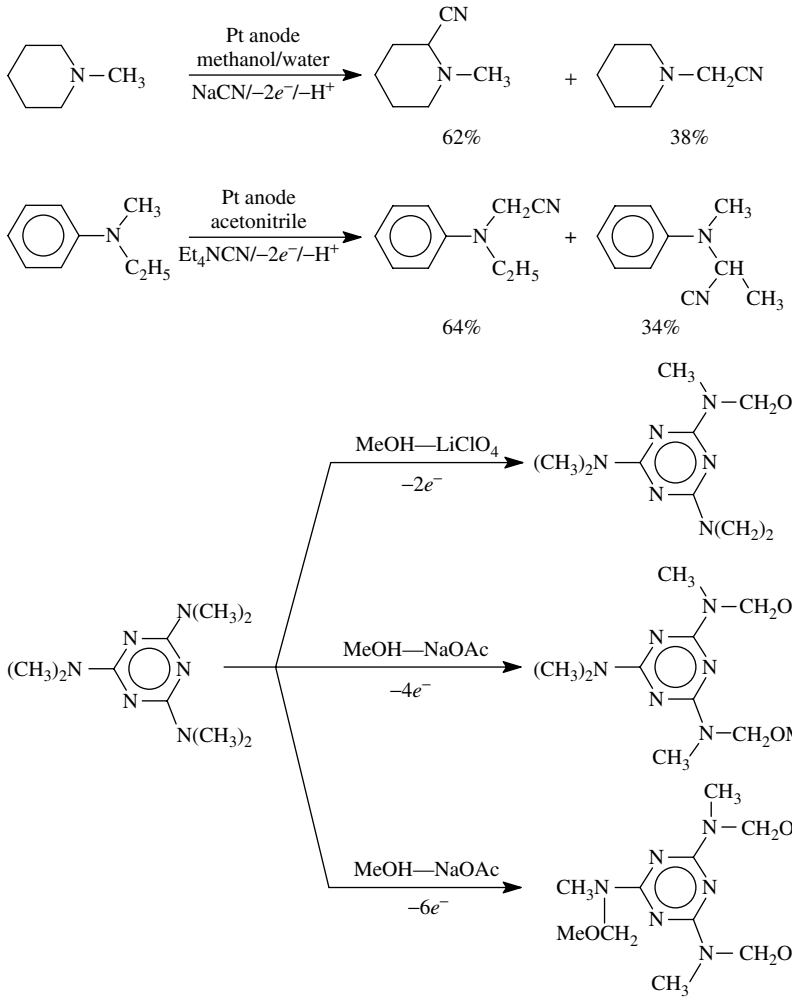


In such reactions, solvent basicity (or its nucleophilicy) may govern the ratio RNu/R(−H). In general, the positive charge induced by the reaction sequence $-e^-/-H^+/-e^-$, which is supposed to occur at the anodic interface, is often borne by a carbon atom. When nucleophilic solvents like H_2O, MeOH, AcOH, acetonitrile, and so on are used, the reaction appears to be (more or less totally) drawn to the substituted RNu compound.

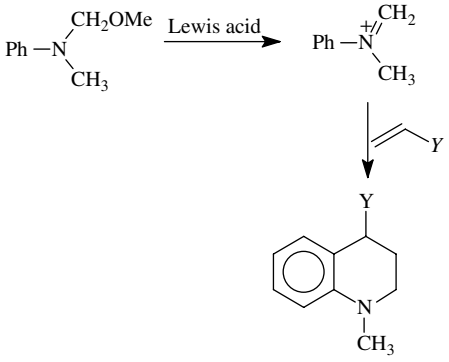
In other cases, a deprotonation reaction occurs. In order to accelerate the loss of the proton, the use of bases is favored, under the conditions that they are weakly nucleophilic (for this purpose, lutidine and collidine are often used in acetonitrile or methylene chloride).

6.3.2.1 Functionalizations of Nitrogen-containing Compounds

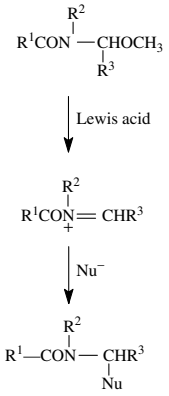
Generally speaking, the electrochemical oxidation of amines yields iminium salts, which are easily hydrolized. However, efficient nucleophiles can be added to the electrolysis mixtures in order to react on the iminium *before* its hydrolysis by residual water. The nucleophiles can be cyanide [36] or methanolate [37]:



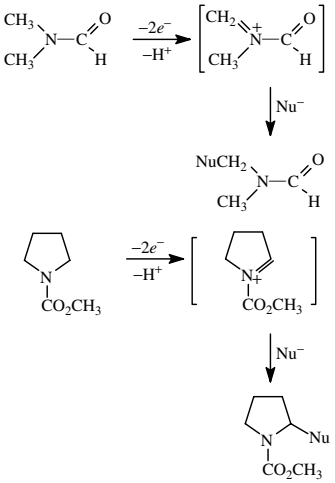
In the example given above, experimental conditions and the control of the electrical charge may orientate the reaction selectively to the mono-, di-, and tri substitution. The action of Lewis acid on substituted amine functions allows one to get back the iminium intermediates [38, 39]. Thus, the nucleophilic action in the course of the electrolysis could be seen as an iminium ''protection'', and renders possible easy synthesis of the heterocycles as outlined below:



Such an iminium transient can be formed and used in an ex situ reaction (off electrolysis) with a very wide palette of nucleophiles. This is particularly true with carbamates [40]:

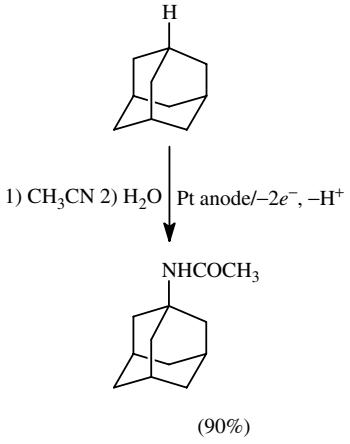


This reaction is achievable with a large palette of nucleophiles such as nitriles, cyanide, R–metal compounds, malonates, $P(OR)_3$, as well as olefins enriched in electrons. Amides (e.g. dimethylformamide) as well as *N*-carboalkoxyamines [41] can also be efficiently substituted by nucleophiles (two-electron oxidations).

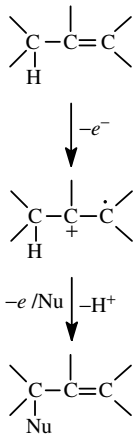


6.3.2.2 Oxidation of C–H Functions

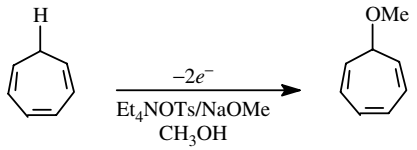
The anodic conversion of tertiary C–H functions (without any specific activation) is possible, yet it needs rather high potentials. One typical example is certainly that of adamantane [42], the oxidation of which can be achieved in acetonitrile. Under these conditions, the corresponding acetamide (via Ritter reaction) was obtained.



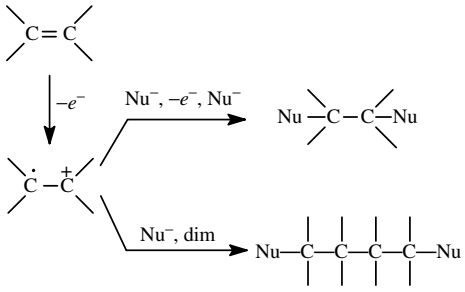
However, the activation of the C–H function by means of a vicinal, electron-rich group (e.g. double or triple bond) strongly increases the reactivity of the system:



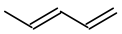
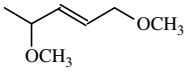
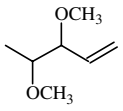
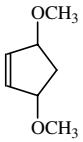
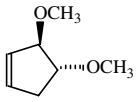
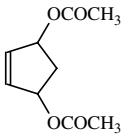
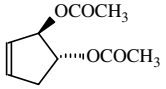
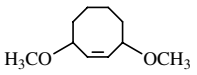
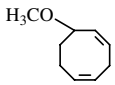
With such reactions, it is quite obvious that deprotonation should be faster than the nucleophilic attack, in order to observe a regioselective substitution. In the equation written above, the ethylenic double bond plays the role of an electrophore (electron transfer from the HOMO of the unsaturated system). The reactivity of unsaturated systems is illustrated by the oxidation of cycloheptatrien [43] in methanol when the methoxylation was achieved in good yield.



N.B. In the absence of a labile hydrogen, the rate of deprotonation at the stage of the radical cation turns out to be so slow that nucleophilic attack on the positively charged carbon of the double bond affected by oxidation becomes the sole reaction to be considered. Thus, alternatively two chemical routes can be envisaged: (1) disubstitution of the (one) double bond and (2) dimerization of the radical cations.

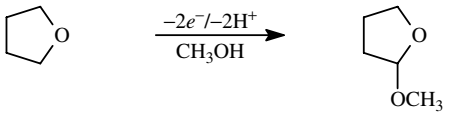


Now, the oxidation of cyclic and acyclic dienes at a glassy carbon electrode quite generally affords 1,4-addition reactions. Side products are also obtained. The table below summarizes [44] the different abilities of double bonds to react anodically (substitution versus addition).

Diene	*Solvent electrolyte*	*Main product*	*Secondary product*
	CH_3OH Et_4NOTs		
	CH_3OH Et_4NOTs		
	CH_3COOH Et_4NOTs		
	CH_3OH Et_4NOTs		

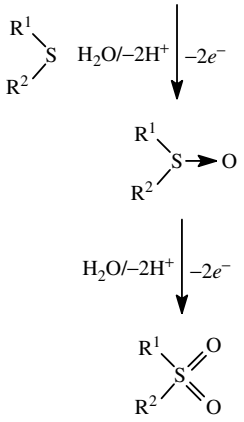
6.3.2.3 **Anodic Functionalization of Ethers**

The anodic functionalization of ethers (tetrahydrofuran, 2-alkyl-1,3-dioxalanes as model molecules) can be easily achieved [45]. Electrolyses are achievable in methanol, mainly at a platinum electrode. Many other nucleophiles are probably candidates to allow other reactions of this type:



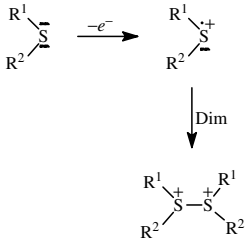
6.4 Anodic Oxidation of Thioethers

Very generally, the anodic oxidation of thioethers in the presence of traces of nucleophiles (e.g. water) affords unavoidably sulfoxides and sulfones. Such double conversion is achieved in two well-separated steps.

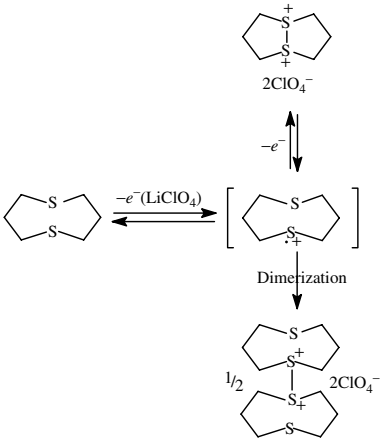


Let us note that until now electrochemistry has not been able to yield chiral sulfoxides from the starting prochiral thioethers even though the use of chiral (chemically modified) electrodes was considered.

On the other hand, in the absence of efficient nucleophiles, the oxidation of thioethers generally creates sulfonium products.

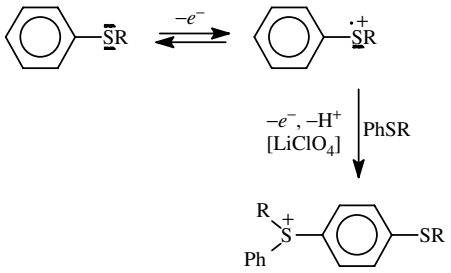


The donating ability of a sulfur atom (soft nucleophile) may contribute [46] to stabilization of $\overset{+}{S}-\overset{+}{S}$ bridges in an intramolecular manner. This is particularly true [47] with cyclic dithioethers, for which transannular interaction plays a role in the product distribution (the one-electron oxidation yields the dimeric dication while the two-electron oxidation yields the monomeric dication). Both forms can be isolated as salts.

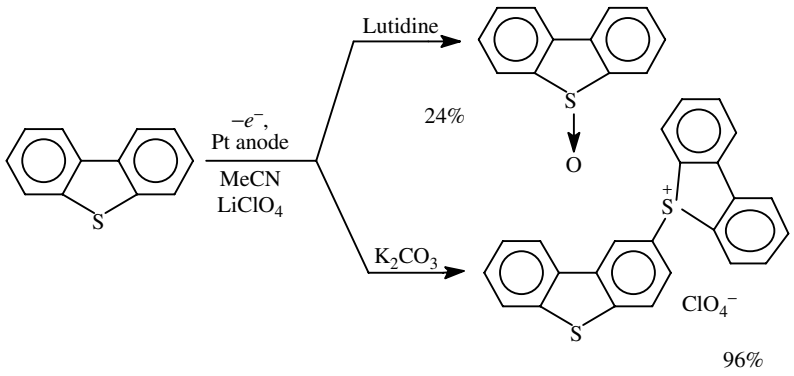


The anodic oxidation of aromatic thioethers allows their conversion [48] to dimeric sulfonium salts via an electrophilic

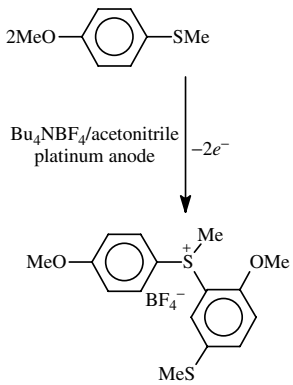
substitution. The reaction is globally monoelectronic:



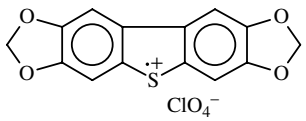
N.B. Competition between the sulfonium formation and that of sulfoxide can be clearly exemplified with dibenzothiophene, where two different types of bases are used in situ (extraction of protons). With lutidine, sulfoxide is obtained probably from the residual water contained in the lithium perchlorate. On the contrary, insoluble bases (like inorganic carbonates) appear to strongly favor a mild electrophilic substitution.



However, such an oxidation process can lead [46] to more complex salts when the para position on the aromatic ring is already occupied.



Lastly, the oxidation of dibenzothiophenes with electron donating substituents, such as alkoxy or amino groups, unexpectedly affords extremely stable radical cations [48] that can be used in organic chemistry as convenient oxidizing species with the oxidative abilities depending on their standard potential (E°). The salt given below is electrically conducting and for this reason is readily obtained in very large amounts (yield: 73%) as a thick deposit on graphite anodes (electrolyte: acetonitrile +0.1 M $LiClO_4$). Many other radicals have been obtained, and their standard potentials have been published [49].



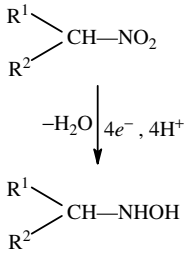
6.5 Electrochemical Conversion of Main Organic Functions in Aqueous and Protic Organic Media

6.5.1 Reduction of Unsaturated Compounds

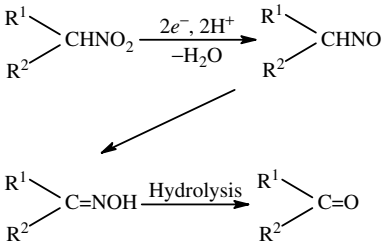
The literature describes mainly reductions in aqueous media using electrodes with high hydrogen overvoltage (mercury pool cathode, lead, graphite). In this review, we chiefly describe the possibility of hydrogenation of derivatives bearing nitro, nitroso, azo, and azoxy functions, as well as oximes and imines, ketones, and aldehydes.

6.5.1.1 Nitro Group

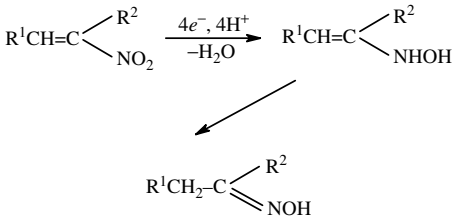
The cathodic reduction of nitroalkanes (primary and secondary) affords [50] (four-electron reduction in acidic or neutral media) the corresponding hydroxylamines that are often quite difficult to be converted into amines.



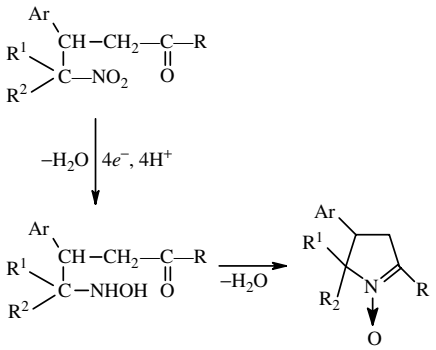
However, in situ rearrangement of two-electron products leads to oximes. Of course, the formation of ketone depends on the stability of a nitroso intermediate in the bulk of the catholyte. The conversion of secondary nitroalkanes into ketones is achieved in high yields. On the contrary, tertiary nitroalkanes cannot give



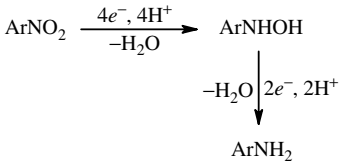
oximes and therefore the regular four-electron reduction allows a selective generation of hydroxylamines. Conjugated nitroalkenes are converted (four-electron reduction) into oximes [51] through the unsaturated hydroxylamine.



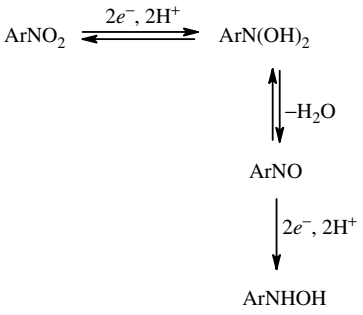
In general, reduced forms of nitroalkanes possessing electrophilic groups (nitro, keto functions) in an adequate position may lead to in situ cyclization reactions [52]. For example:



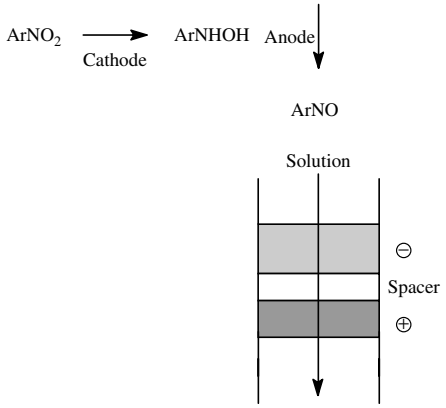
Aromatic nitro compounds ($ArNO_2$), are much more electroactive, and specific reduction steps are seen even in acidic aqueous buffers (e.g. at mercury pool cathodes). In general, two successive steps are observed (a four-electron step followed by a two-electron step) with the formation of hydroxylamine and amine, respectively. The reduction selectivity is therefore very sensitive to the potential applied.



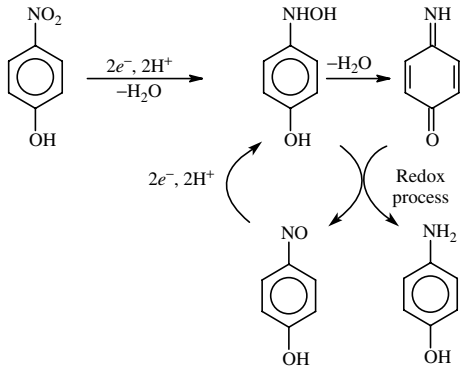
In some cases where the aromatic moiety is substitutionally activated by strongly electron-withdrawing group(s), dihydroxylamine species can be obtained in acidic media [53]. Since



nitroso derivatives are more electroactive than the corresponding hydroxylamines, this method can be useful for producing nitroso compounds far from the electrode; these products can then be trapped in different ways (e.g. a fast reaction with sulfinic acid). In a general manner, nitroso derivatives are produced electrochemically [54, 55] by the use of porous percolating electrodes (carbon felt) of the opposite polarity. With such a cell, hydroxylamine produced at the cathode is immediately oxidized at the associated anode.

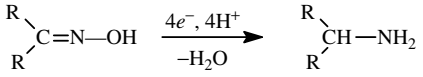


Aromatic hydroxylamines bearing electron-donor substituents (OR, NR_2, SR) exhibit a limited stability. Thus, reduction products are generally the corresponding anilines. The global mechanism is illustrated [56] below for the case of nitrophenol.

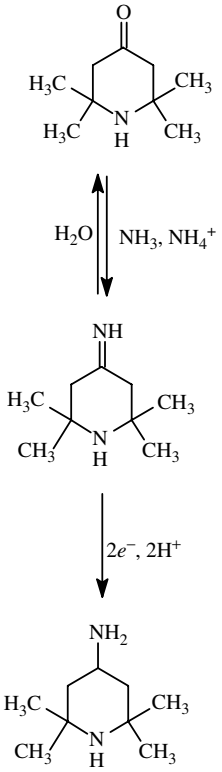


6.5.1.2 Reduction of Functions Possessing at Least One Nitrogen Atom

Oximes are easily reduced [57] in acidic media

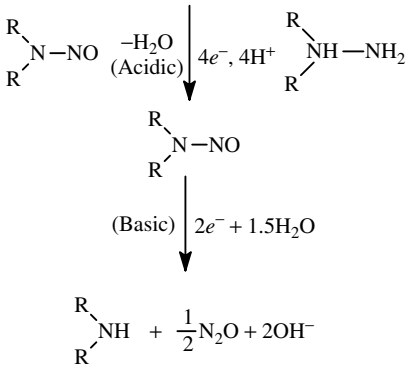


to yield the corresponding amines. The azomethine group is easier to reduce than the corresponding carbonyl group. Imines can be easily formed from ketones in aqueous solutions in the presence of a large excess of ammonia or amines. This method [58] readily allows the cathodic conversion of keto groups to amino functions. Imines (although easily hydrolized in aqueous media) are also reduced into amines according to a two-electron step.

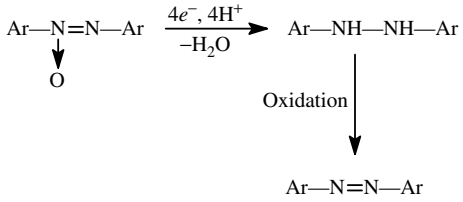


However, Schiff bases (derived from aromatic aldehydes and ketones) lead also to a global two-electron reduction in one or two steps. Reductions at the level of the first step give dimeric diamines.

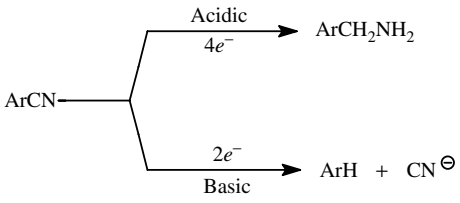
Nitrosamines are converted [57, 59] into hydrazines in acidic media. The product distribution varies dramatically with pH. Thus, in basic media, only secondary amines are obtained.



Azobenzenes and azoxybenzenes are readily reduced to hydrazobenzenes [60]. The reduction of azo compounds was found to be chemically reversible, since the corresponding hydrazo compounds afford easily azo compounds under similar experimental conditions (note that the mercury electrode – easily anodically oxidized – cannot be conveniently used for such experiments). The use of redox cells with both porous electrodes permits the synthesis of azocompounds (a one-pot electrolysis) from azoxy compounds.



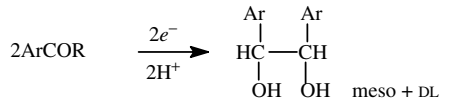
Aliphatic nitriles are not reduced in aqueous solutions. On the contrary, activated aromatic nitriles can be reduced stepwise into amines. The product distribution strongly depends on pH [61]. In acidic media one obtains amines, while neutral or alkaline media lead to cleavage.

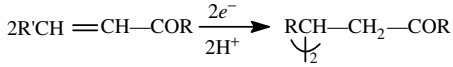


Data concerning the reduction products of a large palette of azo dyes are available [60].

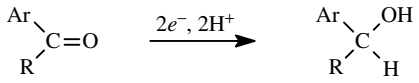
6.5.1.3 Reduction of the Carbonyl Group

In the absence of activation, the carbonyl group (aliphatic ketones and aldehydes) is too weak an electrophore to exhibit a specific reduction in aqueous media. Moreover, aldehydes may lead to hydrates, the reduction of which is even more difficult. On the contrary, activation by a double bond(s) or an aromatic ring in the α position allows one to reduce carbonyl groups. Thus, in acidic media one observes a one-electron step (reduction of the protonated form of carbonyl) that affords the formation of dimers: a mixture of pinacols with aromatic ketones or coupling in ε-diketones with α,β-ethylenic ketones.



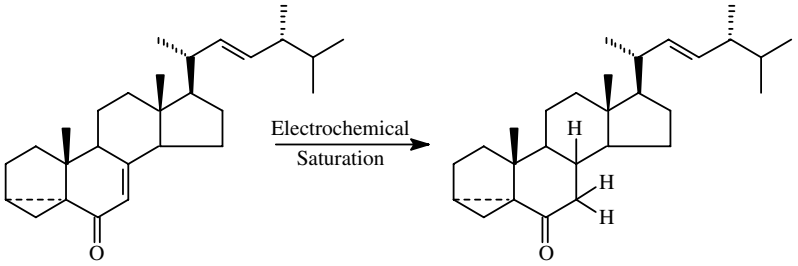


Generally, the use of basic media may direct the reduction process to a two-electron one.

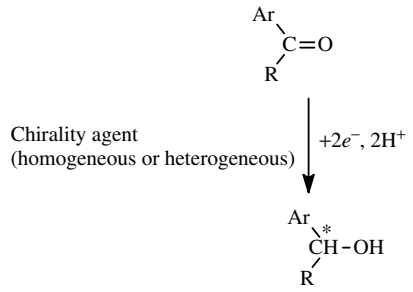


N.B. The use of protic organic solvents (e.g. DMF or acetonitrile containing an excess of a proton donor like water, phenol, benzoic acid, etc.) may be preferred to a hydroorganic solvent owing to the poor solubility of most of organic substrates. An excess of an acidic agent (and its strength) in the catholyte will increase the rate of the protonation and then favor a two-electron reduction. However, a compromise has to be found, because hydrogen evolution may occur in the presence of an excess of acid and mask the specific step of the substrate reduction (poor coulombic yield and possibility the generation of side products by chemical hydrogenation at the cathode interface).

Coupling reactions are not observed in all cases. With highly bulky substrates, radical dimerization becomes so slow that saturation (case of enones) is found to be the main reaction. This is particularly the case with the brassino steroids (potent plant growth promoters), and the compound depicted below for which a stereoselective saturation process was sought. The best conditions for this process are dimethylformamide containing an excess of benzoic acid and a glassy carbon rod as the cathode. The saturated ketone was obtained [62] with a yield of up to 60%. Other proton donors, like trifluoroacetic acid, were used as well.



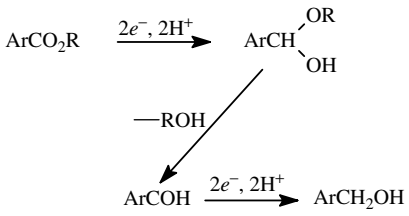
Lastly, the electroreduction of prochiral ketones has been extensively studied, especially by using chiral cathodes or chiral electrolytes.



In most of the cases [63–68], whatever the chosen mode of the chirality induction and the nature of the considered medium, experimental enantiomeric excesses generally are low (ee < 20%) and this technique remains to be optimized.

6.5.1.4 Organic Acids and Derivatives

The electrochemical reduction of carboxylic acids and their associated esters is only possible both in the presence of a strong acid and an efficient activation (Ar accounts for an aromatic ring such as pyridine). The reduction leads successively [69] to aldehyde and alcohol.



Since aldehydes are, in general, more electroactive than the corresponding acids or esters, alcohols can be obtained according to a sole four-electron step. If, on the contrary, the hemiacetal transient is fairly stable, the aldehyde may be isolated from the solution or be continuously trapped as soon as it is formed in the catholyte.

Aromatic amides are also electroactive at mercury or lead cathodes in strongly acidic solutions and afford amines. The use of less acidic solutions [70] leads to the reductive scission of the amide C–N bond.

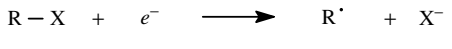
$$\mathrm{ArCONHR} \xrightarrow[-\mathrm{H_2O}]{4e^-,\ 4\mathrm{H^+}} \mathrm{ArCH_2NHR\ (Strongly\ acidic)}$$

$$\mathrm{ArCONHR} \xrightarrow{2e^-,\ 2\mathrm{H^+}} \mathrm{ArCH_2OH + NH_2R\ (Less\ acidic)}$$

6.6 Scission of Chemical Bonds by Means of Cathodic or Anodic Electron Transfers

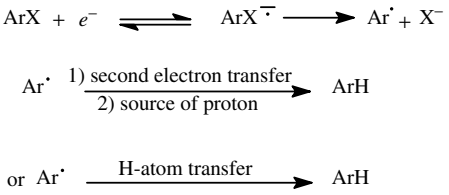
A large number of examples of electrochemical bond scissions (carbon–carbon, carbon–heteroatom, heteroatom–heteroatom) have been reported to date. Quite a large number of these reactions may have applications in organic synthesis, since they may used in the efficient deprotection of various functional groups. One-electron bond cleavages (often through a charged radical) can also be seen as a potential source of free radicals (e.g. alkyl and aryl radicals) and therefore might be fully utilized in radical chemistry (such as the addition to radicalophilic systems and/or induction of radical chains).

Obviously, it is within the reduction range that examples of electrochemical bond cleavages are by far the most numerous. Numerous examples deal with scissions of carbon–halogen, carbon–nitrogen, and carbon–sulfur bonds, but more rarely carbon–oxygen bonds. Mostly such cleavages result [71] in electron transfers to antibonding σ^* or π^* molecular orbitals. In the first case, the uptake of an electron onto the organic substrate and cleavage reaction occur simultaneously: the electron transfer is called *dissociative* (resembling that of a very fast nucleophilic substitution, where the acting nucleophile is the electron itself).



(X: potential leaving group)

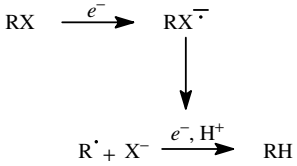
In the second possibility, the electron can be first transferred to a π^* orbital to form a transient radical anion which subsequently evolves the scission sequence. This scheme is often specific to aromatic and heteroaromatic structures.



The electrochemical method is therefore very attractive for generating free radicals in appropriate solvents (not strong H donors). Nevertheless, one has to obtain information about the redox potential concerning the free radical itself (possibly reduced rapidly at the cathode, in the case of a dissociative transfer giving a global two-electron process in which a free radical cannot be detected). In order to circumvent this difficulty, an elegant method is to use redox catalysis to generate free radicals by homogeneous electron transfer from π-acceptor anion radicals (or from any soluble adequate redox system). Also, it is worth noting that when the scission of the Ar–X bond is desired, one should consider the use of aprotic polar solvents.

6.6.1 Main Cathodic Scission Reactions of Carbon–Heteroatom Bonds

Organic halides, especially iodides and bromides, are quite easily cleaved at mercury or glassy carbon electrodes. When not activated, chloroderivatives require highly negative potentials (e.g. −2.8 V versus SCE).

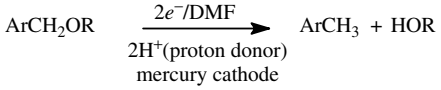


The ease of reduction increases with the number of halogens substituted at a given carbon atom:

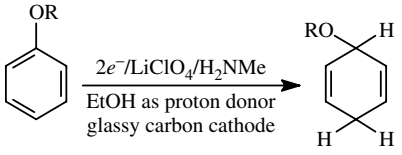


Thus, the persubstitution at a terminal carbon by three fluorine atoms makes possible at least one bond scission (a two-electron reduction).

The position of the heteroatom may also strongly affect the global reactivity of the molecule. Thus, ethers, esters, thioethers, and even amines, that are normally weakly reactive toward the electron, can afford cleavage at the condition to be in an activated position (benzylic or allylic) to the X group (a fast cleavage of the radical anion results in a large shift to less cathodic potentials when the global bielectronic step is considered). In a different way, the reduction of ethers and amines could be achieved only under the conditions of the Birch reaction (i.e. reduction by Li metal in amines or by a solvated electron). These conditions can be electrochemically mimicked (use of amines as solvents and a dried lithium salt as electrolyte). In the two examples given below [1a], it must be recalled that in the first one the cleavage reaction is kinetically prominent, while under the Birch reaction conditions one observes a partial saturation (protonation reaction) of the aromatic ring.

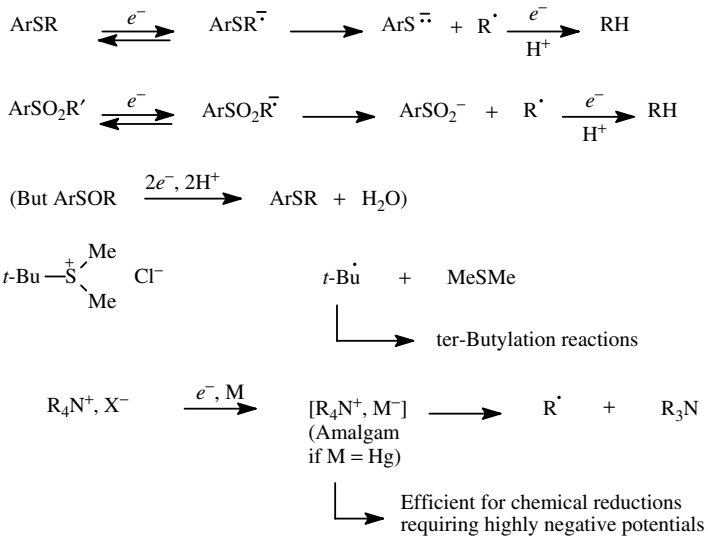


while

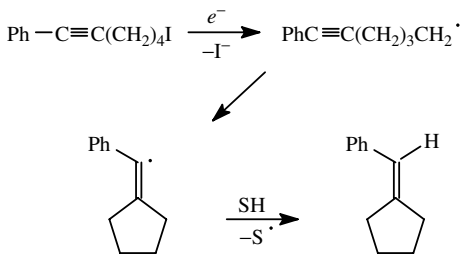


Below are exhibited some useful reactions based on the cathodic cleavage of bonds, which can be seen often as a source of free radicals (homogeneous or heterogeneous cleavage reactions):

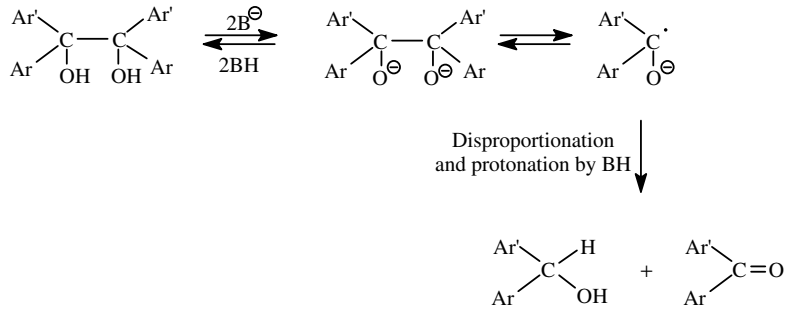
Lastly, intramolecular cyclizations (presumably provoked by the transient radical)



can occur [72]:



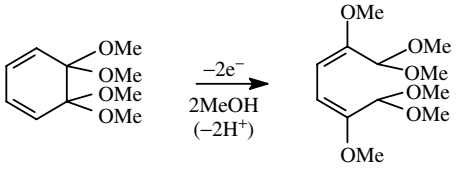
Some studies mention the cleavage of dianion π acceptors, which looks like a homolytic cleavage by cathodic activation. When electron-poor bonds are considered, this type of scission should not be confused [73] with retrodimerization of pinacols (or any dimer with labile hydrogens) in strongly basic media:



6.6.2
Other Types of Bond Scissions

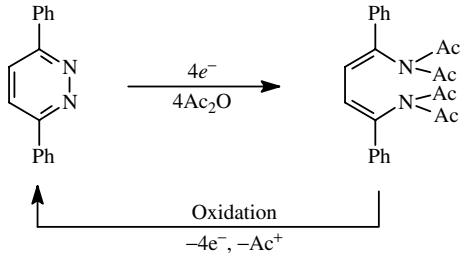
Of course, the facility for scission electron transfer should be influenced by the bond strength and the electron density relative to these bonds. Thus, atom sequences like carbon–carbon and heteroatom–heteroatom (like R—X—X—R) can be cathodically cleaved. The susceptibility of such structures to be cleaved in this manner depend both on the π^* level of the MO relative to R groups and the σ^* level specific to the X—X bond. Many examples can be quoted in the cathodic cleavage of C—C, S—S, S—O, S—N, P—S, P—N, and so on. Some of them are considered to be of economic interest (e.g. cathodic cleavage of the S—S linkage for the conversion cystine into cysteine).

As to the oxidation, an opposite situation may exist [74a] since only strongly electron-enriched bonds are able to undergo scission.



Lastly, let us consider [74b] the case of diphenyl pyridazine, which, when reduced in the presence of acetic anhydride as an electrophile, exhibits a potential reversible closure of the ring in the oxidative

range.



6.6.3 Application of the Electrochemical Scission of Bonds to the Chemical Deprotection

The presence in an organic structure of several functions of similar reactivity toward the same reagent led organic synthesists to the development of the concept of selective protection and deprotection. Numerous protecting groups have been successfully used, but all of them must obey the following criteria:

- the obtaining of the protected substrate has to be selective and is to be achieved in high yields;
- the removal of the protecting group (deprotection process) has to be simple, facile, and without any modifications to the rest of the organic structure; and
- the protecting group must not bring in a supplementary functionality.

However, the major problem met by organic chemists concerns function deprotection, which can be, in some cases, delicate owing to new, unfavorable electronic effects or the necessity to use strongly reactive agents (hydrolysis in highly acidic media or the use of polluting reagents) that are consequently weakly selective.

In many cases, the electrochemical technique recently developed has been shown to be a fruitful alternative to chemical deprotection. Thus, organic chemistry functions that are mainly involved in protection/deprotection processes are mostly aldehydes and ketones, amines, alcohols, and thiols; carboxylic acids can also be quoted, although their esterification as well as their ester hydrolyses are not straightforward.

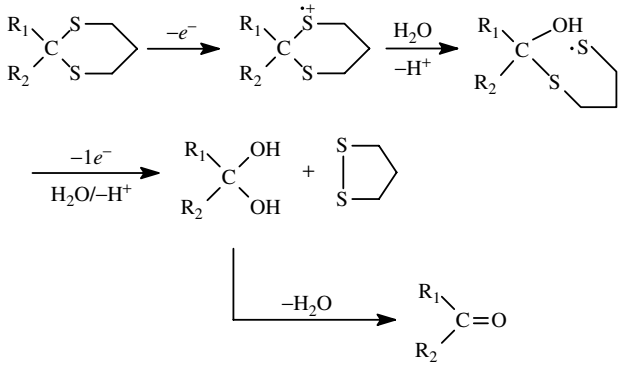
The cleavage of a protecting group by electrochemical means, when such a method is applicable, can be considered as advantageous for the following reasons:

- The equipment is simple (a two-compartment cell with adequate electrodes and a source of current – potentiostat or galvanostat).
- The reaction needs, in principle, only electron transfer. There are therefore no predictable side reactions, but the choice of the solvent (nucleophilic for anodic deprotection) can be crucial.
- The potential or current can be adjusted (and controlled carefully during the course of the reaction) in order to render the deprotection regioselective (control of the potential or the amount of charge).
- Sometimes the use of redox mediators (with electron transfer now taking place in the homogeneous phase) may contribute to better selectivity. Thus, deprotection can be carried out under very mild conditions.

6.6.3.1 Anodic Deprotection of the Carbonyl Group: Cleavage of C—S Bonds

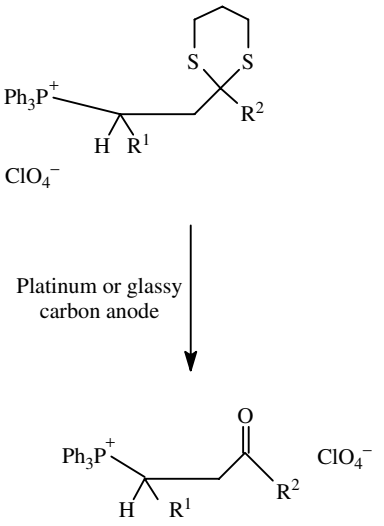
Numerous organic syntheses, particularly those concerning natural products analogs, involve polyfunctional systems

often possessing one or several carbonyl functions. These groups can be easily protected at the stage of acetals, dioxolanes, thioacetals, or dithianes. Their deprotection can be achieved by oxidation processes [75] or hydrolysis reactions [76, 77] (e.g. catalysis by acids or metallic salts, respectively). Especially in the case of dithioacetals and dithianes, the anodic oxidation is now widely considered as a nice method to regenerate carbonyl compounds. Macroscale electrolyses are easily achievable in a several-grams scale allowing the isolation of [78] the wanted aldehyde or ketone (in good yields ranging from 70 to 95%). The experimental conditions are very simple: the oxidation of the sulfur compound is very often carried out at a fixed potential in acetonitrile containing a low percentage of water (1 to 10%) acting as a nucleophile. The electrolyte (concentration 0.1 M) must not be electroactive like $LiClO_4$ or tetraalkylammonium tetrafluoroborate: banish imperatively salts containing halides. In order to avoid progressive poisoning of the platinum anode by sulfur compounds, pulse electrolyses are efficient. Alternatively, glassy carbon electrodes, which avoid generally self-inhibition processes, can be used. The global oxidation, which allows carbonyl group recovery, consumes 2 F mol^{-1} and therefore the current integration can be regarded as an efficient control to obtain maximum yield. The generally admitted mechanism [79, 80] with dithianes involves C−S bond scission at the stage of the radical cation followed by an attack of a nucleophile (usually water). Treatment by acidic water during the workup affords carbonyl function recovery.

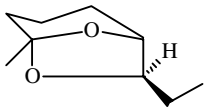


Chemical reactivity of mono- or polysaccharides can be complex owing to the presence of multiple hydroxyl groups, which must be masked temporarily under ester or dioxolane forms. Conditions of deprotection of carbonyl groups of D-galactose, D-mannose, and D-glucose hidden by dithioacetal functions was shown to be successful [81]. The excessive acidity in the anodic compartment of the cell can be efficiently neutralized by the addition of sodium carbonate. Under such conditions, deprotection yields may reach 65 to 85% without the hydrolysis of other protecting groups, if any. Other examples can be given in the field of deprotection of β-ketophosphonium salts [82] with

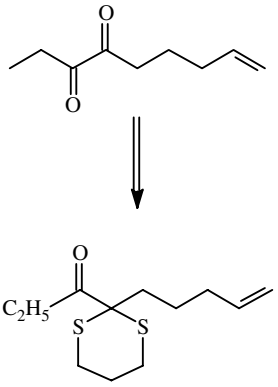
quantitative yields.



The total synthesis of brevicomine, which is the pheromone of a parasite of west American pines



has as a starting building block the following α-diketone chemically obtained as its monodithiane:

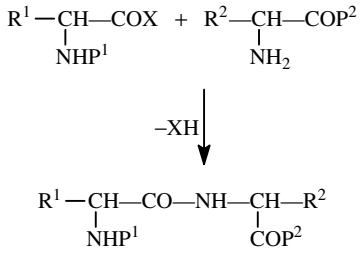


At this level, all chemical deprotection attempts (including concentrated sulfuric acid) failed. On the contrary, the anodic deprotection (potential applied: +1.35 V versus SCE) yielded diketone with a yield of 70%.

Dithianes and gemdithioacetals could be alternatively oxidized indirectly by means of the redox catalysis method. The technique appeared to be particularly mild and mainly avoided inhibition and adsorption phenomena relative to the anode platinum interface. Thus aromatic hydrocarbons (e.g. 9,10-diphenylanthracene) [83] and judiciously substituted triphenylamines [84] afford quite stable cation radicals used homogeneously as oxidants. Their standard potential, E^{o}_{ox}, will determine the rate of electron exchange with the concerned sulfur compound. The cleavage of a C–S bond in any dithiane can be regarded as fast enough to draw the redox catalysis process to the indirect oxidation.

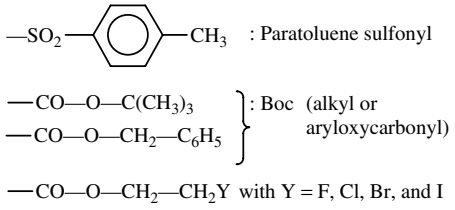
6.6.3.2 Amine Function Deprotection: Cathodic Cleavages of C–S and N–S Bonds

Peptidic synthesis by means of amino acid condensation implies protection of the amine function in one of the substrates and that of the acid function in the other.

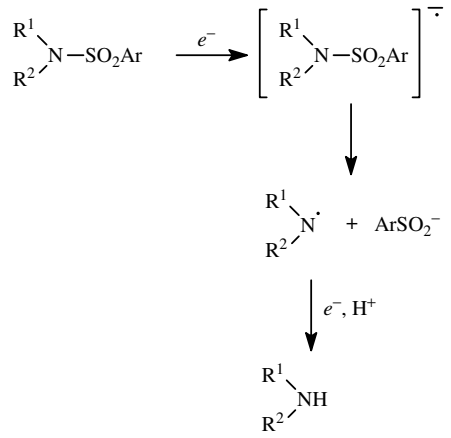


In order to build the amino acid chain, the chemist has to selectively eliminate protecting groups P^1 and P^2.

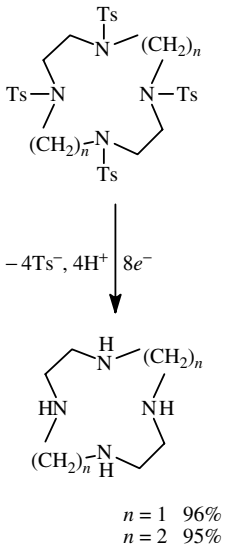
The classical protection modes consist in the replacement of hydrogen atoms by a substituent of strongly withdrawing effect like:



Thus cathodic cleavage of the N—S bond occurs according to a global two-electron process in a nonaqueous medium.

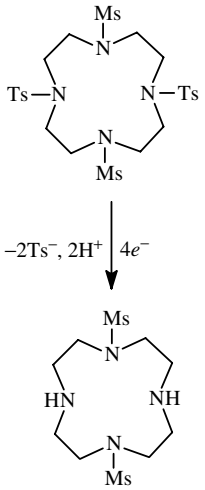


As a matter of fact, the synthesis of polyaminomacrocycles or/and multiheteromacrocycles often needs the step of deprotection of amines blocked under the form of tosyl (Ts) or mesylamide (Ms). The electrochemical method was revealed [85, 86] to be very well adapted to applying deprotection in a very mild and selective manner.

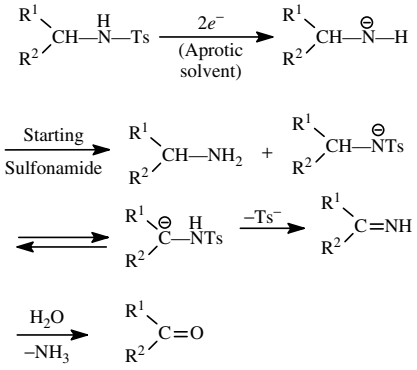


$n = 1$ 96%
$n = 2$ 95%

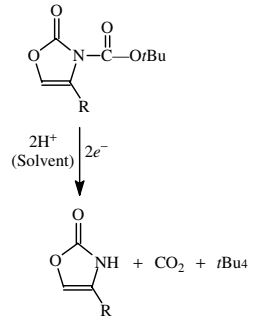
However, only the Ts groups, which are electrophores (π^* level is sufficiently low) in comparison with Ms (mesylate), are electroactive [87, 88]. Thus, the electroactivity difference between these two groups is easily perceived in the reduction of some tetraamides as exemplified below:



However, it is worth noting that the cleavage reaction of secondary tosylamides, as well as of those possessing a benzyl group on the nitrogen atom, must be achieved [89] in the presence of an excess of a proton donor, the pK_a of which is smaller than that of sulfonamide. In fact, the substrate may play the role of a protonating reagent toward EGB (vide infra) leading then to a partial decomposition of the starting tosylamide, more or less, especially of those with benzyl groups.



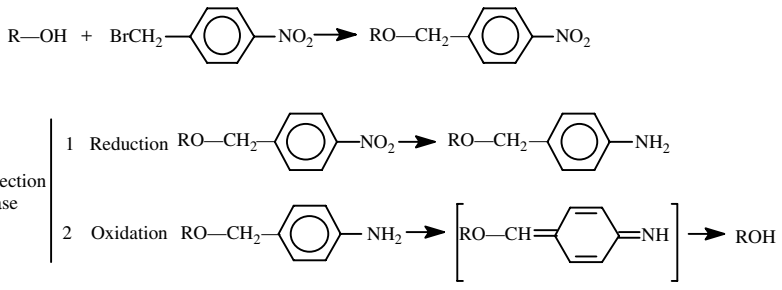
On the other hand, the electrochemical reduction of carbanates (also deprotection of amine functions) occurs [90], but at rather negative potentials (−2.0 V versus Ag^+/Ag system).



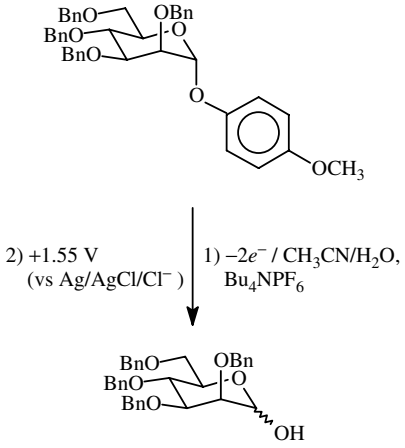
More electroactive protecting groups (such as p-$NO_2C_6H_4SO_2^-$ and p-$NO_2C_6H_4CH_2SO_2^-$) may also be used with amines. Activation by a nitrogroup makes that reduction occur much more easily. Indirect reductions of tosylamide have also been reported.

6.6.3.3 Protection and Deprotection of Alcohols and Thiols

In general, the breaking of the carbon−oxygen bond appears difficult to achieve and needs activation by strongly polar groups. For these reasons, alcohol deprotection is carried out cathodically under the form of benzylic [91] and allylic [92] ethers as well as tosylates [91]. Thus, Torii [93] used the 4-nitrobenzyl group to protect alcohols. The deprotection was carried out in two steps: (1) reduction of the nitrogroup of the amine and (2) oxidation of the amine at a platinum electrode (yield up to 93%).



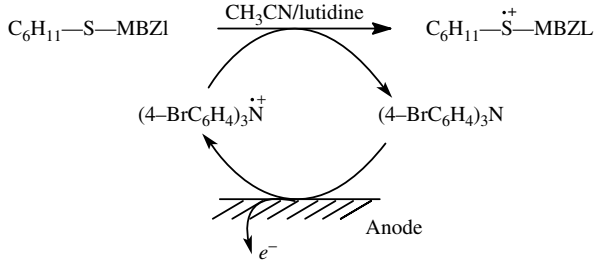
Oligosaccharide synthesis very often needs the protection of the alcoholic functions of a sugar. The specific protection by π-methoxyphenyl and its oxidation giving back the OH function in the anomeric position have been achieved [94] in good yield (74%).

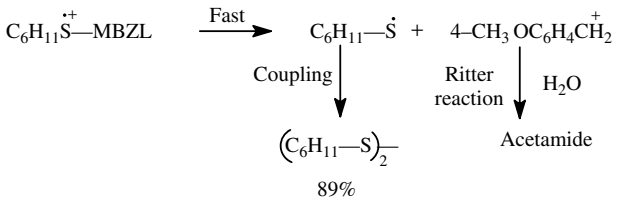


Steckhan [95] used, to good account, the oxidative redox property of 4-methoxybenzyl (MBZL) to achieve thioether deprotection by means of indirect oxidation (thanks to the tris-(4-bromophenyl amine radical cation)). A very similar approach was proposed for alcohol deprotection. The scheme below required that the E^{o} of the triarylamine used as a mediator be lower than E^{o} of the thioethers. The equilibrium is displaced by the fast cleavage reactions at the stage of the radical cation (see below).

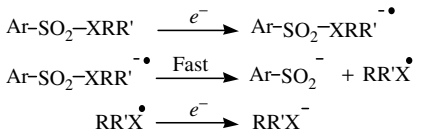
6.7 Electrosynthesis as a Tool for a New Solid-Phase Methodology

Since the 1960s, the tremendous activity devoted to solid phase chemistry has found a renewal of interest with the emergence of a new concept, the combinatorial synthesis [96], which now finds various applications in both academic research and in industry. On the basis of the difference in solubility between a polymeric material and a simple reactant or impurity, the following cleavage from the support and the recovery of the desired product have led to the development of



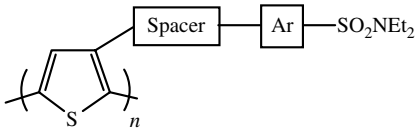


milder and more selective methodologies. Among the wide range of processes used, the photochemical and the electrochemical ones are now evolving. The latter, particularly, involve new conducting polymers electrogenerated at a solid cathode. The originality of this method consists in the cathodic liberation of supported materials. This electrochemical cleavage proceeds through a two-electron mechanism (see Sect. 6.6.3) via a radical anion, which decomposes readily to sulfinate ions and radicals.



Ar : phenyl, naphthyl X : CH, N

The fast protonation in the solution of the resulting anion gives the subsequent product. Taking advantage of this well-known procedure [97], its extension to the design of a new solid-phase methodology has been developed by the Simonet's group. For this purpose, new conducting polythiophenes (PT) bearing an electroactive aryl sulfone or the sulfonamide moiety have been synthesized.

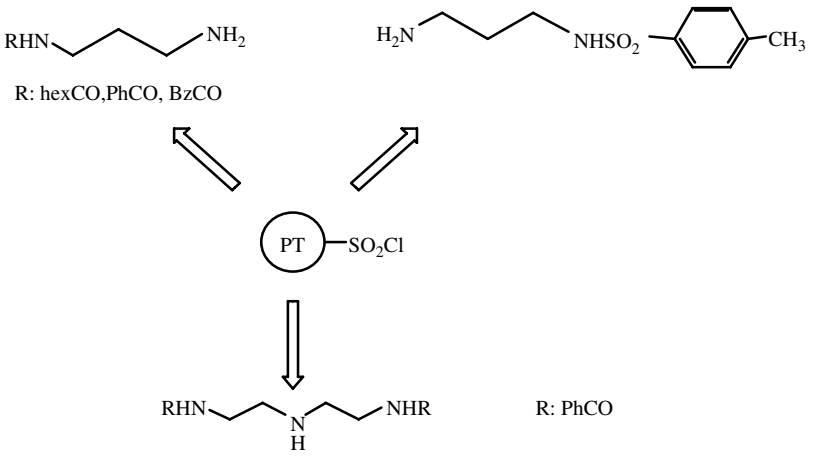


Spacer: $-(CH_2)_2OCH_2-$, $-(CH_2)_2O-$, $-(CH_2)_{1-4}$, $-(CH_2)_2O(CH_2)_2OCH_2-$

Ar : Phenyl, naphthyl

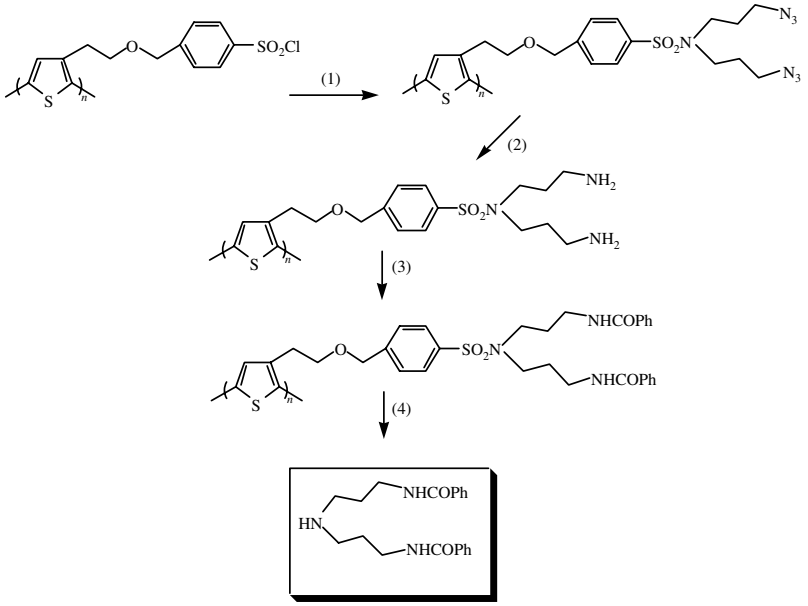
As an example, the PT matrix (spacer:$-(CH_2)_2OCH_2-$and Ar: phenyl) generated by electrodeposition on platinum or on a glassy carbon plate (area = 5 cm^2) at constant potential (+1.40 V versus 0.1 M Ag^+/Ag reference electrode in acetonitrile, anodic charge = 2 C) was found to afford an interfacial functionalization of 2 µmol cm^{-2}. Moreover, a direct functional group accessibility can be electrochemically measured with further complete polymer (PT) recycling by using electroactive probes [98].

Most of the tested, functionalized PTs exhibit satisfactory electrochemical stability with highly reversible *n*- and *p*- doping,



partly because of the presence of a sulfonamide moiety [99]. This new interfacial process, starting from the electrogenerated sulfonylchloride PT, has been validated in the synthesis of various polyamine precursors, which were obtained with yields ranging from 40 to 60%.

Thus, the deprotected secondary amine shown in the scheme below has been obtained in 60% yield according to the following sequence:



(1) $[N_3(CH_2)_3]_2NH$, pyridine; (2) $SnCl_2$, PhSH, Et_3N; (3) PhCOCl, pyridine; (4) $2e^-$, proton source

Besides their good electrochemical stability, most of the tested polymers present an identical behavior toward usual chemical reagents as well as a functionalization similar to the ''classical'' polystyrene beads. All supported molecules formed by successive chemical steps were cleaved off from the conducting resin by the cathodic S−N bond scission of the aryl sulfonamide moiety.

This electrochemical cleavage appeared to be faster (less than 1 mn) than the parent chemical [100] or photochemical [101] processes. These mild and fast conditions should therefore be of particular interest for the synthesis of fragile compounds such as acid labile derivatives. Finally, considering the small amount of products obtained at the cathode interface, this methodology currently seems fully applicable for combinatorial chemistry.

6.8 What is the State of the Art for Organic Silicon Derivatives?

6.8.1 Electroreduction

The electroreduction of silicon derivatives has to be focused on organylchlorosilanes, which usually consist of a Si−Cl bond cleavage resulting in silyl anionic

species. Because this process requires highly negative potentials, not many organic compounds are electrochemically inactive under these conditions to be suitable electrophiles in the reaction with electrogenerated silyl anions. More often, it is a starting molecule itself that acts as a trap for silyl anions, so the reduction of chlorosilanes results in Si−Si products.

Main products of the electroreduction of organylchlorosilanes in carefully dried organic polar solvents are the corresponding hexaorganyldisilanes [102].

$$2R_3SiCl \xrightarrow[-2Cl^-]{+2e^-} R_3SiSiR_3$$

When reducing a mixture of two monochlorosilanes, cross-disilanes were obtained, usually along with some amount of a symmetrical disilane formed from the easiest-to-reduce starting chlorosilane [103].

$$R_3SiCl + Me_3SiCl \xrightarrow[-2Cl^-]{+2e^-} R_3SiSiMe_3$$

R = Alk, Ar

Similarly, reducing mono- and dichlorosilanes and disilanes taken in different combinations (e.g. $R_2SiCl_2 + R'_3SiCl$; $R_3SiSiR'_2Cl + R_3SiCl$; $2R_3SiSiR'_2Cl$; $2R_3SiSiR'_2Cl + R_2SiCl_2$, etc.) yield tri-, tetra-, and pentasilanes [102, 104, 105].

Silyl anions, electrogenerated by the cathodic reduction of polyfunctional chlorosilanes, react with remaining Si−Cl bonds of the starting molecules or of intermediately formed oligomers to give linear polysilanes. The molecular mass of the resulting polysilanes depends on the electrolysis conditions (solvent, electrode material, supporting salt, etc.) [102, 105–108].

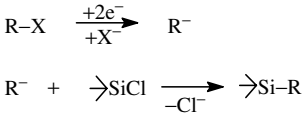
$$R_1R_2SiCl_2 \xrightarrow[-2Cl^-]{4e^- \text{ to } 8e^-} \left(\begin{array}{c} R_1 \\ | \\ Si \\ | \\ R_2 \end{array}\right)_n$$

R_1, R_2 = Ar, Me

Since organylchlorosilanes are good electrophiles in S_N2 reactions, another possibility for cathodic synthesis of organosilicon compounds is the reaction of an electrogenerated nucleophile with various chlorosilanes. The most developed are processes involving C−nucleophiles, obtained by the electroreduction of an appropriate halogen derivative, more often a chloride. Some of these processes have turned out to be efficient alternative synthetic routes to C−Si compounds and witnessed the award of several patents [109]. The active center of such a C−nucleophile can be C_{sp^3} [107] or C_{sp^2} [110] carbon. It seems that C_{sp} carbanions, electrogenerated by reducing a C−H bond in phenylacetylene, can also react with chlorosilanes, though the mechanism of formation of silylated products is not yet clear [111, 112]. In this reaction, linear and cyclic alkylsilanes [107, 109, 113], allyl and vinylsilanes [114], benzylsilanes [115], and aryl [110, 116], and heteroarylsilanes [117] were obtained. The use of polyhalogenated aromatics in this process provides an easy and selective route to polysilylated aromatic products [110].

Not only halogenorganic derivatives but also various organoelement compounds, which are easier to reduce than chlorosilanes, can be involved in a similar process. A number of electrogenerated organoelement nucleophiles substitute Cl in chlorosilanes to give different silylated

products at relatively low cathodic potentials.

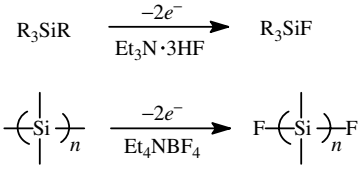


Here R^- represents N^- [118], O^- [119], S^- [120], Se^- [121], Ge^- [122], and Sn^- [123] nucleophiles. The electrochemical generation of nucleophiles from the substrates initially inactive toward chlorosilanes, affected immediately in the presence of the latter, provides a very convenient way for obtaining a large palette of organosilicon products.

Finally, let us note that the electroreduction of trimethylarylsilanes themselves leads to silylated products with the loss of aromaticity [124].

6.8.2 Electrooxidation

Among the most thoroughly explored reactions in anodic silicon electrochemistry are the processes of direct and indirect anodic substitution, both based on the high ability of the Me_3Si group to cationic elimination. In the first group of these reactions, a silicon substrate undergoes oxidation giving electrophilic species that react with an external nucleophile [125]; in the second, an organic compound is oxidized to produce an electrophile that attacks organosilanes, usually allylsilanes, β-silylated electron-rich systems [126,] or silyl enol ethers [127]. Thus, a large variety of processes of direct anodic substitution were realized, providing convenient ways for functionalization of allyl, benzyl, carbonyl, carboxy, ether, amino, and thio derivatives. An example of such a reaction is the anodic fluorination of silanes and polysilanes [125]. Below, reactions using two different types of fluorinating reagents are shown:



Oxidizing tetraorganylsilanes in the presence of some carboxylic acids, alcohols, or water yield the corresponding benzyl and allyl derivatives [126, 128]:



R = Ph, CH_2CH;

Nu = ROH, RCOOH, H_2O.

A similar process with β-silyl amines, carbamates [129], ethers [130], sulfides [130], or selenides [131] also results in the displacement of the Me_3Si group with an appropriate nucleophile.

Indirect anodic substitution was used for a functionalization of heteroatom (O, S, N) α-substituted organylstannanes and organylsilanes, and of triphenylphosphine [127].

6.9 Some Words about the Use of Selenium in Organic Synthesis

6.9.1 Electroreduction

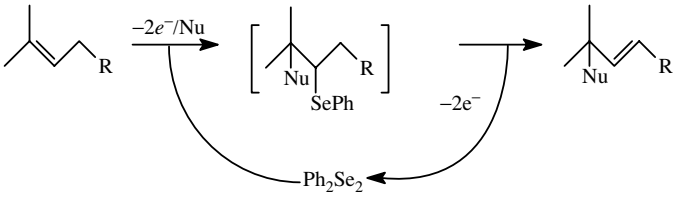
The behavior of selenium organic compounds appears to be similar to that of the sulfur analogs. The electrochemical reduction of a mixture of Ph_2Se_2 with arylhalogenides bearing acceptor substituents [132], or with an aliphatic

bromide [133] results in corresponding arylselenides.

$$Ph_2Se_2 \xrightarrow[-X^-]{+2e^-,\ 2RX} 2\ RSePh$$

The use of sacrificial Se electrodes, producing Se^{2-} and $Se_2{}^{2-}$ anions under cathodic polarization, allows one to obtain diorganylselenids and diselenides, respectively [134].

Likewise, in organic chemistry, the β-elimination of selenoxides is very important when electrochemically oxidizing organylselenides containing a $(Se)C_{sp^3}-H$ fragment. This reaction, combined with the above selenylation, is a key step in a whole series of selenium-catalyzed processes of the functionalization of organic compounds [138a, 139].

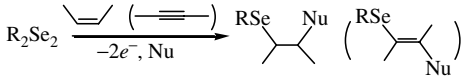


$$2Se^- \xrightarrow[2RX,\ -2X^-]{+2e^-} R_2Se_2$$

6.9.2 Electrooxidation

High localization of the positive charge on the Se atom in electrochemically generated species often results in many interesting processes, both possibly involving a cleavage of bonds formed by Se.

First reported by Torii [135] and Bewick [136], anodic selenenylation of alkenes or enol compounds gives rise to various RSe adducts bearing a nucleophile at a vicinal position [137].



If the unsaturated compound contains an internal electron donor center that is able to act as a nucleophile in this process, intramolecular etherification or lactonization is observed [138].

Electrooxidation of diphenylselenide [140] and of a number of cyclic aryl selenides (having no acidic protons at the α-carbon prone to syn elimination), potentially interesting from the biological point of view, gives selenoxides in different yields [141].

High acidity of a proton in the β position to selenium, increasing even more when the molecule obtains the positive charge upon oxidation, can be used for its oxidative replacement. Thus, the reactions of α-methoxylation, acetoxylation, and fluorination (Nu = MeO, AcO, F) of alkylarylselenides [142] have been realized.

$$PhSeCH_2R \xrightarrow[Nu,\ -H^+]{-2e^-} PhSeCH(Nu)R$$

Another interesting process is the electrosynthesis of triarylselenonium salts, hard to obtain through traditional chemical ways [143].

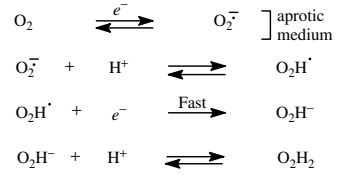
$$Ph_2Se_2 \xrightarrow[ArH,\ -H^+]{-2e^-} Ph_2\overset{+}{S}eAr$$

6.10 Electrochemical Activation of Small Inorganic Molecules Usable in Organic Synthesis

Small molecules can be advantageously activated by electron transfer at cathodic or anodic interfaces. Thus the reduction of dioxygen into the superoxide $O_2^{\cdot-}$ as the primary reaction appears now as being very trivial. But carbon dioxide, sulfur dioxide, tellurium, selenium, and sulfur can be cathodically activated as well, with the formation of nucleophilic entities, the reaction of which in situ can be considered. On the other hand, electrophiles can be generated from nitrogen nitroxide or molecular sulfur. Below are detailed (vide infra) reactions that are related essentially to dioxygen and sulfur, in order to achieve organic synthesis of molecules. Conception and building of electrochemical sensors in relation to the detection of small electroactive molecules do not find their place in the present chapter. Lastly, the anodic oxidation of halides and the electrochemical halogenation of organic molecules (addition, substitution) are treated separately (see section 6.11).

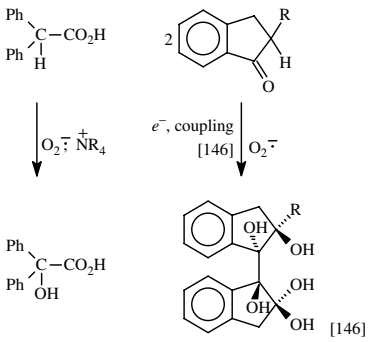
6.10.1 Cathodic Intermediates from Dioxygen and Principal Types of Reactions

Potassium superoxide KO_2 is commercially available. But since dioxygen is reversibly reduced in aprotic organic solvents at many types of electrodes (e.g. platinum, glassy carbon), it can be easily formed in situ [144]. The superoxide can be readily protonated: therefore aprotic solvents possessing an efficient proton donor and aqueous solvents permit a two-electron step to occur. This step affords the formation of hydrogen peroxide. Its further reduction can be achieved (conversion into water), but at very negative potential values. A large number of studies exist concerning the electrocatalysis of the latter conversion at chemically modified electrodes (e.g. at electrodes covered by polymers containing metalloporphyrin sites).



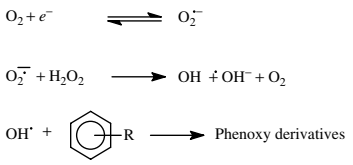
The use of molecular oxygen in electroorganic synthesis must take into account the fact that superoxide can act both as a reducing reagent and nucleophile. Additionally, in the presence of fairly acidic substrates (but *not* electroactive at the standard potential of O_2 (in acetonitrile $E_{O2} = -1.06$ V versus SCE)) acidobasic reactions may take place in the bulk of the catholyte (owing to the basicity of $O_2^{\bullet-}$ and O_2H^-). Below are described some standard reactions employing dioxygen (or its reduced forms) generated in DMF, principally in the presence of tetraalkylammonium salts.

- acidic substrates (carbon substitution) [145 146]

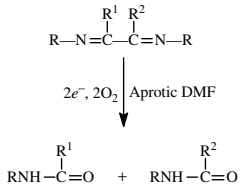


[146]

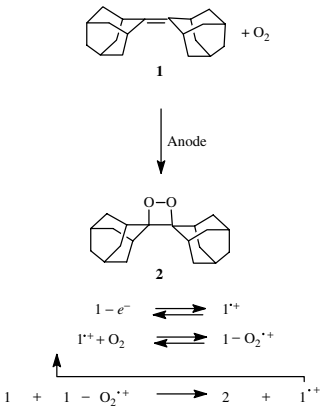
- generation of the hydroxyl radical from hydrogen peroxide [147] (electrochemical Fenton reaction)



- coupling superoxide/radical anion of the π–acceptor. Cleavage of C–C bonds [148]



N.B. Dioxygen can be implicated in reactions induced by anodic electron transfers. It is particularly the case in the oxidation of electron-rich double bonds in the presence of molecular oxygen. The reaction below exemplifies such a catalytic process (here in the case of adamantylideneadamantane) with the formation of dioxetane [149, 150]. *Note that in this case the reaction is catalytic in electron (i.e. catalytic amount of charge extracted by the anode).*

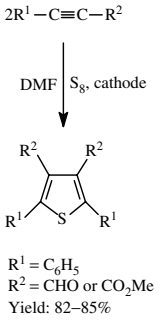


6.10.2 Reduction of Carbon Dioxide

CO_2 reduction at metallic electrodes is generally poorly selective [151]. Monoelectronic reduction of carbon dioxide may occur at a platinum cathode in nonaqueous solvents, but at very negative potentials. Catalytic activation of CO_2 has been described (e.g. at a cathode modified by a rhenium complex in a hydroorganic solvent): the observed conversions did correspond to the formation of CO and formic acid. In organic synthesis, CO_2 was mainly used as an electrophile (toward electrogenerated anions from π–acceptors or electrogenerated nucleophiles when adequate transition metals ions were present in situ) for the purpose of carboxylation.

6.10.3 Cathodic Behavior of Sulfur, Tellurium, and Selenium: an Easy Way of Forming Specifically Reactive Nucleophiles

Although sulfur (S_8) is poorly soluble in organic solvents, particularly in polar ones, it can be electrochemically activated by the use [152]of solid sulfur/carbon electrodes (produced by melting and solidifying mixtures S/C generally in the ratio 2/1 w/w). Under electrochemical polarization, nucleophiles from S_8 can be formed and allowed to react in situ with nonelectroactive electrophiles. Product distributions after the electrolyses are in agreement with the formation of nucleophiles principally like S_2^- and S_2^{2-}. Thus the use of sulfur sacrificial cathodes permitted elegant synthesis of persubstituted thiophenes [153]:



Many other examples of the addition of electrogenerated nucleophiles from sulfur on activated olefins are available [154]. However, such reactions are rarely specific and regioselective.

Similarly, the cathodic formation of Se_2^{2-} and Se^{2-} (to be reacted with a large palette of electrophiles in situ) from solid selenium (with the help of ultrasound for activating its reduction) has also been reported. This method appears to be of interest for a safer chemistry when using organic derivatives of selenium and tellurium.

6.10.4 Anodic Activation of Molecular Sulfur and of S—S Bonds

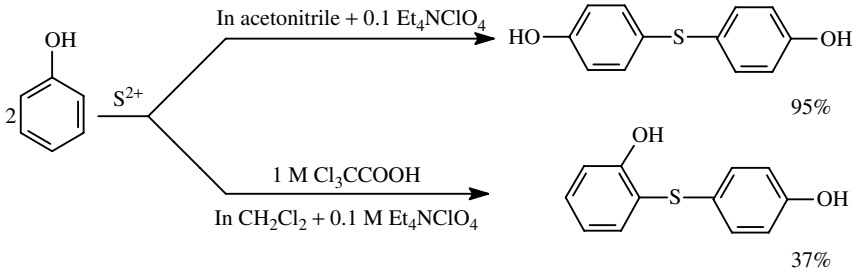
The use of sulfur sacrificial anodes readily allows the activation of sulfur, principally in methylene chloride. The reaction takes place at rather positive potentials (at a platinum anode) and was shown to correspond to the equation:

$$S_8 \xrightarrow[\text{Tetraethylammonium tetrafluoroborate/}CH_2Cl_2\text{ Pt anode}]{+2.0\text{ V vs SCE}} 8S^{2+} + 16e^-$$

After electrolysis, the solution of $[S^{2+}, 2BF_4^-]$ in CH_2Cl_2 is transferred into a flask (the absence of moisture is necessary) and is allowed to react with a large palette of nucleophilic entities or donors. Thus, with aromatic systems, C—S—C linkages can easily be formed.

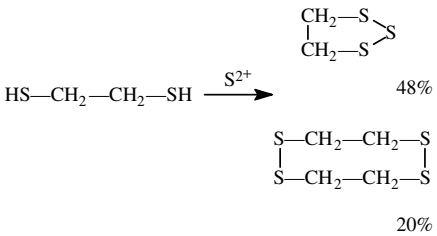
$$2XArH + S^{2+} \longrightarrow XAr—S—ArX + 2H^+$$

The regioselectivity [155] can be changed with the acidity of the considered medium.

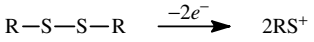


With other nucleophiles, such as thiols [156], S—S—S linkages are generated in good yield. However, the side formation of S—S bonds could be provoked by the oxidative properties of S^{2+}.

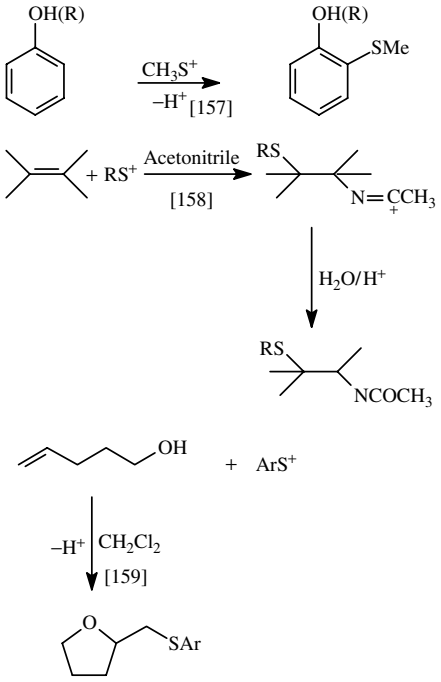
$$2RSH + S^{2+} \longrightarrow RS—S—SR + 2H^+$$



N.B. Similarly, useful electrophiles can be readily generated from the anodic oxidation of disulfides.

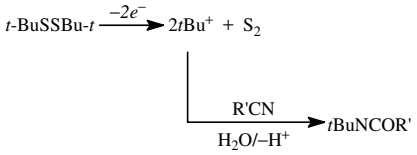


Through this path, electrophilic substitution and addition can be achieved on aromatic and unsaturated substrates. It appears to be a good way to introduce the MeS group in a fairly regioselective manner.



[157–159]

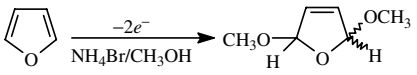
However, the above schemes do not work with a tertiary R. The oxidation of the corresponding disulfides yields essentially tertiary carbocation. Consequently, oxidations performed in nitriles [160] can be considered as an elegant route (via the Ritter reaction) to a large palette of amides.



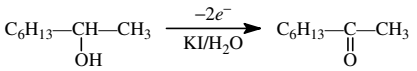
6.11 Oxidation of Halides

The anodic oxidation of I^-, Br^-, and Cl^-, and the active species generated by the oxidation may be useful in organic synthesis. The use of halides in electrochemistry has to be classified in two categories: (1) generation of the oxidizing species and (2) halogenating reagents.

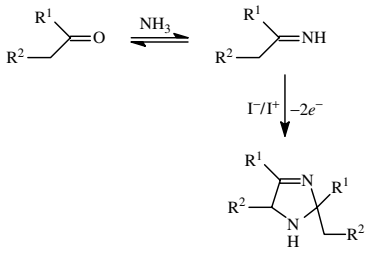
Thus, Br^- and I^- can advantageously be used as mediators in some specific, indirect oxidations. Let us quote the oxidation of furan [161] in methanol in the presence of a catalytic amount of ammonium bromide:



as well as the oxidation of alcohols [162,] in excellent yield, in the presence of potassium iodide.

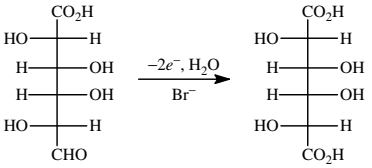


Such a mediator was successfully used [163] to obtain 2,5-dihydro-1*H*-imidazoles from ketones in liquid ammonia.

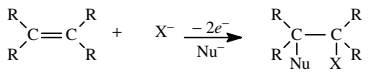


Br^- and Cl^- were also employed to prepare, from C=C bonds, epoxides in excellent yields [164].

Such a process was also recently applied for the successful industrial synthesis of mucic acid from the indirect oxidation of galacturonic acid [165].



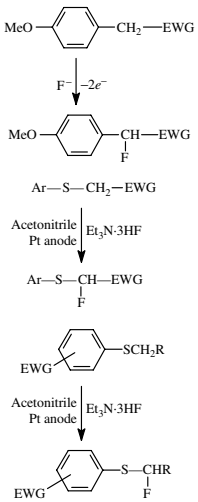
On the other hand, the generation of I_2, Br_2, and Cl_2 at the anodic interface may afford specific reactions toward organic substrates. Thus, unsaturated bonds that are electroinactive can lead [166] to addition, especially in a nucleophilic solvent like acetonitrile.



Accordingly, the iodoamidation of cyclohexene in acetonitrile was shown to result in quantitative yields. The chlorination of benzene derivatives [167] as well as the iodination of anisole (mixture of the three isomers [168]) was also reported. However, these substitution reactions required very positive potentials, suggesting that the organic substrate has to be concomitantly oxidized.

6.12 Anodic Fluorination

The introduction (by substitution or addition) of one or several fluorine atoms to an organic molecule can be achieved anodically [169]. For that several methods or processes exist. Electrolyses can be conducted directly in anhydrous hydrogen fluoride (Simons process). Such anodic reactions (e.g. substitution of alkanes) assume in principle the primary oxidation of the substrate (into a radical cation or a cation) with which the oxidation products are allowed to react in situ with HF. Thus, such fluorination reactions have to be achieved at very positive potentials and are often poorly selective. On the other hand, fluorination can be also achieved in the presence of the fluoride ion in conventional solvents like acetonitrile, nitromethane, and sulfolane. In such cases, the F^- ions (as nucleophiles) belong to the electrolyte: use principally of R_4NF, $R_4NF{\cdot}xHF$, or $Et_3Nx{\cdot}HF$. Numerous substitution reactions have been reported, especially when the CH function is activated by the presence of a electron-withdrawing group in the α-position. For example:

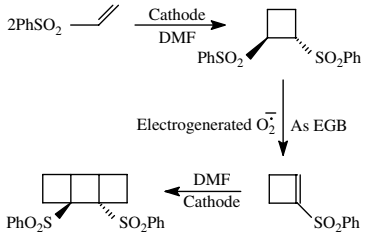


At more positive potentials, another fluorination step occurs and allows also the formation of gem-difluorinated compounds in high yields.

6.13 Zero-current Electrolyses

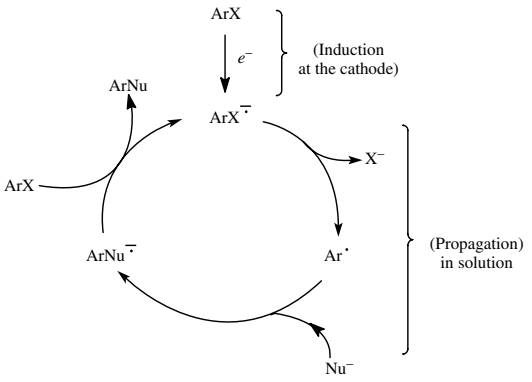
Electrochemistry has also been widely used to form EGBs in the cathodic compartment and EGAs in the anodic compartment in order to carry out specific reactions. Such bases and/or acids can be formed from probases and/or proacids, judiciously chosen in order to obtain the pKs necessary to induce organic reactions in situ. Also, in this review, one can quote some examples of the formation of free radicals capable of inducing radical chains. In this respect, the use of electrochemistry for triggering of S^1_{RN} reactions [170, 171] via electrogenerated transient species ($\overrightarrow{E}\,C\,\overleftarrow{E}$ processes, in which the electricity consumption is practically zero) is worth mentioning. This reaction was mainly studied in the case where an electrophilic aryl radical Ar$^\bullet$ is formed (via heterogeneous electron transfer, essentially). The main conditions are: (1) a fast addition of the nucleophile onto the aryl radical, (2) a fast electron exchange between $ArNu^{\cdot-}$ and ArX, and (3) the use of solvents that are H donors, in order to limit the formation of ArH, which is considered as a termination reaction of the radical chain.

Lastly, cyclodimerization and cycloadditions can also be achieved thanks to ion radicals possibly produced via anodic or cathodic means. Let us quote the cyclodimerization of vinylphenylsulfone (well known as a dienophile in Diels–Alder additions), which can afford in good yield a totally new dienophile [172] that is particularly useful in cyclobutane compound chemistry. Its cyclodimerization which is catalyzed by electron transfer leads to ladderanes as well [173]:



6.14 Elements of Electrochemical Chiral Synthesis

To a certain extend, electrochemistry can be successful [1] in chiral organic

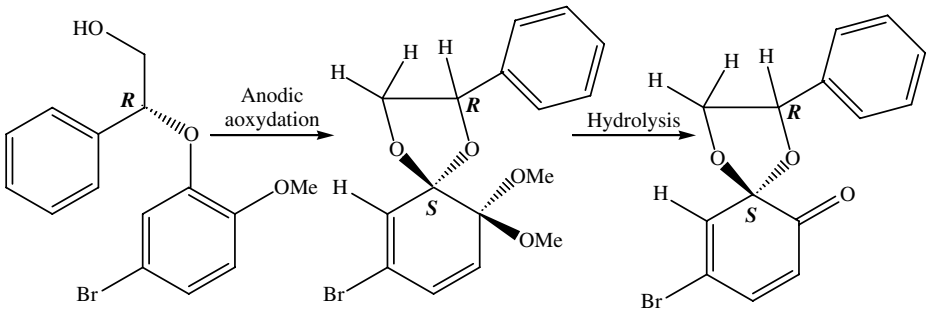


synthesis. Many methods have been used, such as the building of chiral interfaces (modified electrodes by chiral electrode absorbants [174]) or with specific polymer-coated electrodes [175], the achievement of asymmetric synthesis thanks to chiral media [176] or electrolytes [177], as well as the development of electrogenerated chiral oxidants (such as Os-catalyzed asymmetric dihydroxylation of olefins with Sharpless's ligand [178]).

Recent works in asymmetric induction has yielded two new ways: (1) the electroenzymatic reduction [179, 180] and (2) the use of oreganometallic complexes as mediators (like complexes of Ni^{II}, Pd^{II}, or Co^{II} [181, 182]. For example, nickel-catalyzed cross coupling between aryl halides and α-chloropropionic acid derivatives bearing chiral auxiliaries affords arylpropionates in high yield and excellent enantiomeric excess [183].

On the other hand, the induction bound to the presence of chiral auxiliary was elegantly applied [184] both within the cathodic and anodic ranges. Below is described a recent example in which the anodic oxidation allowed the conversion of arylmethyl ethers bearing a chiral ethanol unit into different dimethyl ketal units.

The reaction is stereoselective, and by simply varying the attachment point of the chiral pro-metal auxiliary to the starting arene, all four possible monoketal diastereomers were shown to be accessible.



Only one diastereoisomer !

6.15 Some Models of Electrolysis Cells: From Laboratory Cells to Scaling up

The present chapter gives a long, but far-from-exhaustive, list of reactions achievable by means of electrochemistry at cathodes or anodes under specific experimental conditions (choice of the solvent, the electrolyte, the material of the working electrode, etc.). However, the geometry of the cell, strictly depending on the distance and the volume between the working electrode and the counterelectrode, and the nature of the separator, if any, can play a fundamental role in the success of the electrochemical method for achieving a synthesis.

Below are presented a few cells, some made of glass, of small volume, which could be convenient for organic chemists to test on small amounts of substrates (e.g. from 0.1 to 10 g), their possible interest, and the feasibility of electrochemical conversion. Some other cells usable for

industrial applications are quoted for their principle.

6.15.1 Laboratory Cells

The simplest cell ever designed is certainly that used for achieving the oxidation of carboxylates: one beaker with two large area electrodes (the working electrode made of a grid of platinum) dipping in a solution of sodium methanolate in methanol. There is no separator, as there is no need for the same under an inert atmosphere. The cell is connected to a simple battery.

In general, if a precise electrochemical potential of the working electrode is necessary (in order to avoid side reactions), a three-electrode cell appears necessary. The control of the electrochemical potential of the working electrode is ensured thanks to a potentiostat. Fig. 1 shows a small cell (a four-fingered cell) working with about 4 to 6 ml of liquid electrolyte, a simple separator, and a nitrogen bubbling system. The amount of substrate submitted to electrochemical conversion could be about 10^{-4} mol. Another possibility for use with larger amounts of the electroactive substrate is the H-shaped cell with a separator (glass frit) between the two compartments of the cell (Fig. 2). For a better distribution of potential at the working electrode surface, cylindrical cells are preferred: here the cathode and the anode (of the same area) are placed at parallel positions on either side of the separator. Figure 3 gives an example of a convenient cell of this type, which is easy to conceive and is excellent for collecting gaseous effluents.

Especially in the field of anodic oxidations (conveniently adapted to platinum or glassy carbon electrodes), concentric capillary-gap cells are useful with solvents of low dielectric constant such as methylene chloride. Such cells do not have separators, and the reaction at the counter electrode (and products) should not interfere

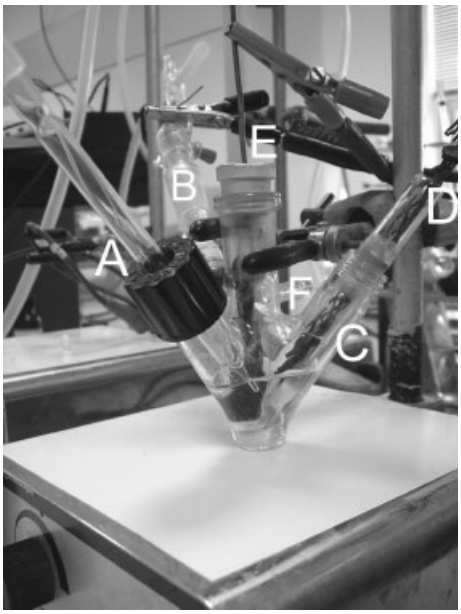


Fig. 1 View of a glass cell (capacity: 5 to 7 ml of solution) connected to a power source (potentiostat). Case of a reduction at a copper cathode. A: inlet for inert gas, B: reference electrode like a saturated calomel electrode, C: anodic compartment (a glass tube ended by a glass frit), D: platinum grid as anode, E: working electrode: copper grid (area: about 4 cm^2), F: inert gas outlet. The solution is stirred with a magnetic bar. Thus for an amount of electroactive compound (one-electron reduction) of 10^{-3} mol, electrolysis current could be of the order of 0.1 A and the reaction completed (until nil current) in much less than 1 hour.

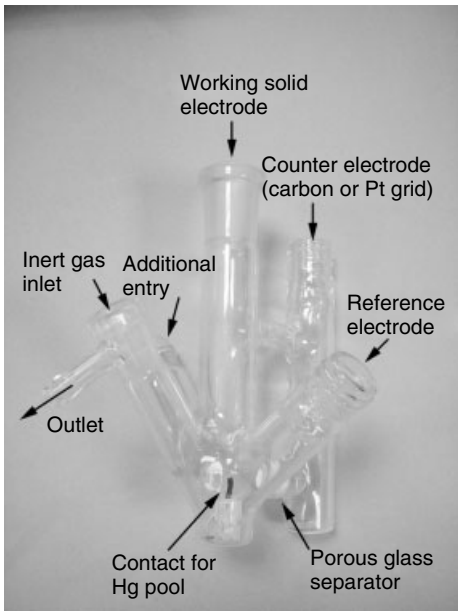


Fig. 2 Glass vessel used as H-shaped cell. Capacity of the compartment: 10 to 20 ml of liquid electrolyte. A solid electrode (cylindrical shape) can be used but this type of cell fits well with a small size (diameter: 2 cm) mercury pool.

with the starting material and/or the reaction product formed at the working electrode. Most of the time, the cathodic reaction leads to hydrogen evolution.

Lastly, cells equipped of porous electrodes are useful in organic electrochemistry. Redox flow cells were efficiently developed by Moinet [55], and this concept can be considered as a great advantage toward conventional reduction and oxidation methods in organic chemistry. It is essentially because they can be particularly useful to obtain electroactive intermediates in reduction and oxidation multistep processes. A striking example had been mentioned in Sect. 6.5.1.1 in the synthesis of aryl nitroso compounds. However, the flow of the solution through the two porous electrodes is difficult to regulate.

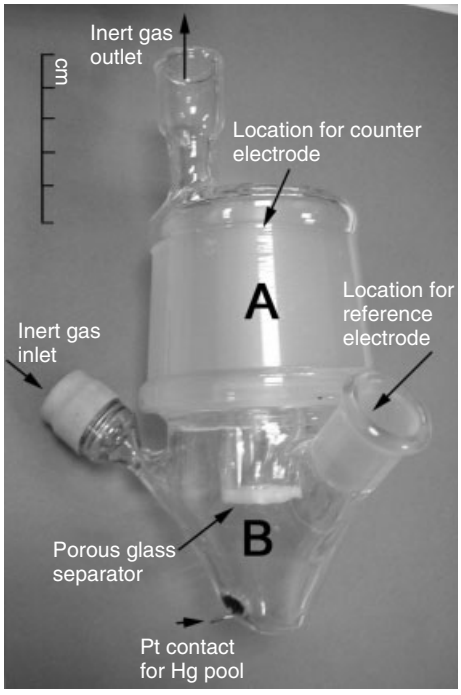


Fig. 3 Glass vessel (two compartments) having a large volume (50–100 ml) working compartment. Its main interest is to provide a short distance between anode and cathode and therefore a better distribution of current lines.

6.15.2 Industrial Cells

Within the last three decades, a large number of versatile electrochemical cells have been developed [185]. Beside their suitable geometry and principle of operation of the cell, the availability of more stable electrodes and separators (mostly membranes of acceptable lifetime) has considerably improved.

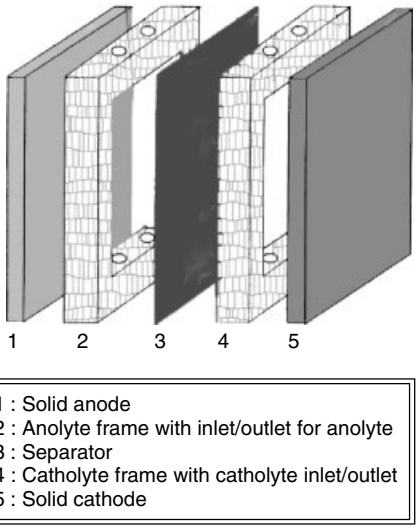


Fig. 4 Schematic representation of a plate-and-frame cell (according to [185]).

The ideal cell in order to scale up an electrochemical reaction can depend on the reaction, the electroactivity of the substrate to convert, the concentration of the substrate, as well as the current density at the working electrode. The use of a separator is necessary when the electrode can affect the whole process negatively. With anodic oxidations, the reaction at the counter electrode is most frequently the cathodic formation of hydrogen. In these cases, a separator does not seem indispensable: a tank cell (kind of Grignard type reactor equipped with cylindrical electrodes) or a capillary-gap cell (piling of bipolar electrodes in a cylinder-shaped vessel connected to an anodes and a cathode located at the top and the bottom of the cell) can be considered as suitable devices for anodic conversions. More generally, the so-called plate-and-frame cells (Fig. 4) are used in a battery.

6.16 Possible Industrial Applications of Organic Electrosynthesis

It is well known that electrochemistry is the main source for the synthesis of

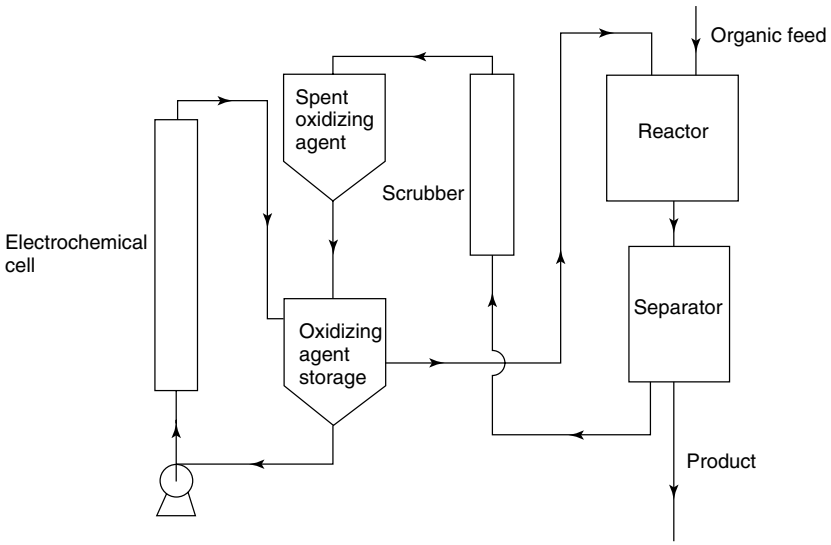


Fig. 5 Flowsheet for an indirect electrochemical oxidation process with a separate chemical (from [186]).

reagents that can be (also) reducing (like alkali metals) or oxidizing (like transition metals ions of high valency) species. More specifically, direct processes can be scaled up by means of indirect (called *ex-cell processes*) or direct electrolyses. By ''indirect'' we mean that the electrolysis cell is only a conventional device capable of forming continuously a reducing or oxidizing reagent (possibly of short lifetime) (Fig. 5) transferred into a reactor where the redox reaction can occur.

It is difficult to draw up on a yearly basis an exact list of industrial processes using the electrochemical technique. The situation evolves all the time, principally depending on the cost of energy in each country and the will of companies to invest in new processes. The scale also depends on the cost of chemical conversion and the specificity of the electrochemical method compared to the more conventional organic chemistry routes. We give below a table that summarizes *some* electrogenerated commercial products (according to Refs 186–188)).

Most of the products quoted in the table above derive from electrochemical reactions given in the chapter. Obviously, the most accessible experimental conditions (like cheap and nonpolluting electrodes like graphite and steel, aqueous solvents or methanol, nonbuffered media) are preferred for reasons of cost and environmental requirements.

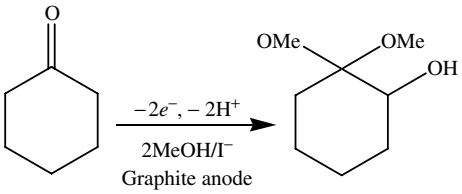
Thus, acrylonitrile (an example of activated olefin) yields cathodically a coupling reaction (adiponitrile) in aqueous solution (use of quaternary salts) containing tetraethylammonium *p*-toluene sulfonate, which provides a hydrophobic thin layer at the cathode interface [189, 190]. This process first designed with a divided cell was transformed to

	Product (cathodic (C) or anodic (A))	*Precursor*	*Company/Country*	*Scale (metric tons per year)*
C	Adiponitrile	Acrylonitrile	Monsanto (inventor)	300 000 in 1998
C	2,2-Dichloro-1-(4-hydroxyphenyl) ethene	3,3,3-Trichloro-1-hydroxy-1-(4-hydroxyphenyl) ethene	Japan	120
C	3-Hydroxybenzyl alcohol	3-Hydroxy benzoic acid	Mitsui Taatsu	100
C	Tetrafluoro-*p*-xylene	Bis-quaternary ammonium salt	Europe	200
A	Dimethyl sebacate	Monomethyl adipate	Asahi, other	?
A	Anisic aldehyde	4-Methoxytoluene	BASF	3500
A	Tolualdehyde	*p*-Xylene	BASF	1000
A	Anthraquinone	Anthracene	L.B. Holliday	100
A	2,5-Dimethoxy-2,5 dihydrofuran	Furan	BASF	100
A	Calcium gluconate	Glucose	India, Sandoz	>1000
A	Fluorocarbons	Carboxylic acids	3 M, Bayer, Phillips	>10
C	4-Aminophenol	Nitrophenol	India	?
C	Azoxybenzene	Nitrobenzene	India	3
A	Mucic acid	Galacturonic acid	EDF	?
A	Acetoin	Butanone	BASF	?

fit to an undivided cell. This example shows the versatility of the electrochemical method.

Electrolyses in aqueous solutions were applied in saturation reactions (partial hydrogenation of phthalic acid in a divided cell at a lead cathode [191]) and oxidation of carboxylates by the Kolbe reaction (coupling leading to sebacic acid in methanol/sodium methanolate at a platinum anode [192]).

A successful anodic substitution in the α-position of electron rich group was used [193] for realizing the oxidation of cyclohexanone. Such a process is achieved in methanol in the presence of KI. The electrogeneration of "I^+" as a mediator in the process is believed to facilitate the whole process.



A valuable goal appears to be the anodic conversion of substituted toluenes into the corresponding aldehydes. The reaction can be achieved either in methanol [194] (intermediate formation of a ketal) or in aqueous solution in an indirect manner (presence of Mm^{II} or/and Ce^{III} ions as mediators [195]). The indirect oxidation of polyaromatic hydrocarbons (naphthalene, anthracene) into the corresponding quinones could be achieved in the presence of electrogenerated ceric ions.

The selective cleavage of carbon–chlorine bonds is also of interest for industrial purposes (e.g. cathodic mono dehalogenation of dichloroacetic acid [196]). Similarly, the reductive scission of C−Cl bonds may lead under more delicate conditions (use of an aprotic organic solvent like dimethylformamide, undivided cell, soluble anode process) to stabilized anions capable of reacting in situ with CO_2 (synthesis of fenoprofen according to the process of SNPE and EDF in France [197]).

References

1. (a) H. Lund, O. Hammerich, *Organic Electrochemistry*, 4th ed., Marcel Dekker, New York, 2001; (b) T. Shono, *Electroorganic Synthesis*, Academic Press, London, 1991; (c) J. Grimshaw, *Electrochemical Reactions and Mechanisms in Organic Chemistry*, Elsevier, 2000.
2. J. H. P. Utley, R. J. Holman, *Electrochim. Acta* **1976**, *21*, 987–989.
3. R. N. Renaud, P. J. Champagne, S. Savard, *Can. J. Chem.* **1979**, *57*, 2627.
4. L. Becking, H. J. Schäfer, *Tetrahedron Lett.* **1988**, *29*, 2797–2800.
5. G. Le Guillanton, J. Simonet, *Tetrahedron Lett.* **1990**, *31*, 3149–3150.
6. J. Simonet, unpublished results.
7. K. Nyberg, *Acta Chem. Scand.* **1971**, *25*, 534–542.
8. L. Eberson, K. Nuberg, H. Sternerup, *Acta Chem. Scand.* **1973**, *27*, 1679–1683.
9. J. M. Chapuzet, J. Simonet, *Tetrahedron* **1991**, *47*, 791–798.
10. V. Le Berre, L. Angely, N. Simonet-Gueguen et al., *New J. Chem.* **1989**, *13*, 131–137.
11. V. Le Berre, L. Angely, N. Simonet-Gueguen et al., *New J. Chem.* **1985**, *9*, 419–426.
12. Y. Rollin, M. Troupel, O. G. Tuck et al., *J. Organomet. Chem.* **1986**, *303*, 131–137.
13. S. Torii, H. Tanaka, K. Morisaki, *Tetrahedron Lett.* **1984**, *26*, 1655–1658.
14. A. Jutand, S. Negri, A. Mosleh, *J. Chem. Soc., Chem. Commun.* **1992**, *23*, 1729–1730.
15. C. Amatore, A. Jutand, L. Thouin et al., *L'actualité Chim.* **1998**, *8–9*, 43–62 and references cited therein.
16. M. Durandetti, M. Devaud, J. Périchon, *New J. Chem.* **1996**, *20*, 659.
17. C. Amatore, A. Jutand, *Organomeallics* **1988**, *7*, 2203–2214.

18. C. Amatore, in *Novel Trends in Electroorganic Synthesis* (Ed.: S. Torii), Kodansha, Tokyo, 1995, pp. 227–230.
19. M. Durandetti, S. Sibille, J. Y. Nédélec, in *Novel Trends in Electroorganic Synthesis* (Ed.: S. Torii), Kodansha, Tokyo, 1995, pp. 209–212.
20. M. Durandetti, S. Sibille, J. Y. Nédélec et al., *Synth. Commun.* **1994**, *2*, 145–146.
21. M. Durandetti, J. Y. Nédélec, J. Périchon, *J. Org. Chem.* **1996**, *61*, 1748–1755.
22. C. Gosmini, J. Y. Nédélec, J. Périchon, *Tetrahedron Lett.* **1997**, *38*, 1941–1942.
23. S. Condon-Gueugnot, E. Léonel, J. Y. Nédélec et al., *J. Org. Chem.* **1995**, *60*, 7684–7686.
24. R. Scheffold, S. Abrecht, R. Orlinski et al., *Pure Appl. Chem.* **1987**, *59*, 363–372.
25. S. Condon-Gueugnot, J. Y. Nédélec, *Novel Trends in Electroorganic Synthesis* (Ed.: S. Torii), Springer-Verlag, Tokyo, 1998, pp. 387–390.
26. S. Ozaki, H. Matsushita, H. Ohmori, *J. Chem. Soc., Perkin Trans.* **1993**, *19*, 2339–2344.
27. M. Troupel, Y. Rollin, J. Périchon et al., *New J. Chem.* **1981**, *5*, 621–625.
28. J. F. Fauvarque, C. Chevrot, A. Jutand et al., *J. Organomet. Chem.* **1984**, *264*, 273–281.
29. (a) B. Giese, *Radicals in Organic Synthesis: Formation of Carbon-Carbon Bonds*, Pergamon Press, Oxford, NY, Seoul, Tokyo, 1986; (b) G. Stork, in *Selectivity – A Goal for Synthetic Efficiency* (Ed.: W. Bartmann, B. M. Trost), Verlag Chemie, Basel, 1984.
30. (a) R. D. Little, M. K. Schwaebe, in *Topics in Current Chemistry* (Ed.: E. Steckhan), Springer-Verlag, Berlin, 1997, pp. 1–48, Vol. 185. (b) K. D. Moeller, Schwaebe, in *Topics in Current Chemistry* (Ed.: E. Steckhan), Springer-Verlag, Berlin, 1997, pp. 49–86. (c) J.-Y. Nédélec, J. Périchon, M. Troupel, Schwaebe, in *Topics in Current Chemistry* (Ed.: E. Steckhan), Springer-Verlag, Berlin, 1997, pp. 141–173.
31. W. J. M. Van Tilborg, C. J. Smit, *Recl. J. R. Netherl. Chem. Soc.* **1978**, *97*, 89–90.
32. T. Shono, *Electroorganic Synthesis*, Academic Press, London, 1991, p. 98.
33. R. Scheffold, M. Dike, S. Dike et al., *J. Am. Chem. Soc.* **1980**, *102*, 3642–3644.
34. T. Shono, *Electroorganic Synthesis*, Academic Press, London, 1991, p. 103.
35. T. Shono, *Electroorganic Synthesis*, Academic Press, London, 1991, p. 102.
36. T. Chiba, Y. Takata, *J. Org. Chem.* **1977**, *42*, 2973–2977.
37. M. Cariou, K. El. Hobbi, J. Simonet, *Electrochim. Acta* **1993**, *38*, 2481–2484.
38. N. L. Weinberg, E. A. Brown, *J. Org. Chem.* **1966**, *31*, 4058–4061.
39. T. Shono, Y. Matsumura, K. Inoue et al., *J. Am. Chem. Soc.* **1982**, *104*, 5753–5757.
40. T. Shono, *Electroorganic Synthesis*, Academic Press, London, 1991, p. 71.
41. T. Shono, Y. Matsumura, K. Tsubata, *Org. Synth.* **1985**, *63*, 206–213.
42. A. Bewick, J. M. Mellor, B. S. Pons, *J. Chem. Soc., Chem. Commun.* **1978**, *17*, 738–740.
43. T. Shono, T. Nazoe, H. Maekawa et al., *Tetrahedron Lett.* **1988**, *29*, 555–556.
44. T. Shono, A. Ikeda, *Chem. Lett.* **1976**, 311.
45. T. Shono, Y. Matsumura, O. Onomura et al., *Synthesis* **1987**, *3*, 1099.
46. M. N. Elinson, J. Simonet, *J. Electroanal. Chem.* **1993**, *350*, 117–132.
47. G. S. Wilson, D. D. Swanson, J. T. Klug et al., *J. Am. Chem. Soc.* **1979**, *101*, 1040–1041.
48. K. Uneyama, S. Torii, *Tetrahedron Lett.* **1971**, *12*, 329–332.
49. (a) M. Cariou, T. Douadi, J. Simonet, *New J. Chem.* **1995**, *19*, 65–76. (b) M. Cariou, T. Douadi, J. Simonet, *Bull. Soc. Chim. Fr.* **1996**, *133*, 597–610.
50. P. E. Iversen, H. Lund, *Acta Chem. Scand.* **1966**, *19*, 2303–2328.
51. P. E. Iversen, H. Lund, *Tetrahedron Lett.* **1967**, 4027–4030.
52. (a) M. Cariou, R. Hazard, M. Jubault et al., *J. Chem. Res. (S)* **1986**, 184–185. (b) M. Cariou, R. Hazard, M. Jubault, et al., *Can. J. Chem.* **1983**, *61*, 2359–2366.
53. A. Darchen, C. Moinet, *J. Electroanal. Chem.* **1975**, *61*, 373–375; **1976**, *68*, 173–180; **1977**, *78*, 81–88; *J. Chem. Soc., Chem. Commun.* **1976**, 487–488.
54. C. Lamoureux, C. Moinet, A. Tallec, *Electrochim. Acta* **1986**, *31*, 1–12.
55. C. Lamoureux, C. Moinet, *Bull. Soc. Chim. Fr.* **1988**, 59–65.
56. A. Darchen, D. Peltier, *Bull. Soc. Chim. Fr.* **1972**, 40–4081; **1972**, 4061–4067; **1973**, 1608–1611.
57. H. Lund, *Acta Chem. Scand.* **1959**, *13*, 249–267.

58. H. Lund, O. Hammerich, *Organic Electrochemistry*, 4th ed., Marcel Dekker, New York, 2001, pp. 466–467.
59. G. Jacob, C. Moinet, A. Tallec, *Electrochim. Acta* **1982**, *27*, 1417–1423.
60. J. Simonet, N. Gueguen-Simonet, in *The Chemistry of the Hydrazo, Azo and Azoxy Groups* (Ed.: S. Patai), John Wiley & Sons, Chichester, 1997, pp. 391–463.
61. J. Volke, V. Skala, *J. Electroanal. Chem.* **1972**, *36*, 383–388.
62. C. Brosa, C. Rodriguez-Santamaria, J. F. Pilard et al., *Phys. Chem. Chem. Phys.* **2001**, *3*, 2655–2661.
63. L. Horner, D. Degner, **1971**, 1245–1248; *Electrochim. Acta* **1974**, *19*, 611–627.
64. R. N. Gourley, J. Grimshaw, P. G. Millar, *J. Chem. Soc., Chem. Commun.* **1967**, *3*, 1278–1279; *J. Chem. Soc. (C)* **1970**, 2318–2323.
65. (a) M. Jubault, A. Lebouc, A. Tallec, *Electrochim. Acta* **1982**, *27*, 1339–1341; (b) M. Jubault, E. Raoult, D. Peltier, *J. Chem. Soc., Chem. Commun.* **1979**, 232–233.
66. D. Seebach, H. A. Oei, *Angew. Chem.* **1975**, *87*, 629–630.
67. B. F. Watkins, J. R. Behling, E. Kariv et al., *J. Am. Chem. Soc.* **1975**, *97*, 3549–3550.
68. B. E. Firth, L. L. Miller, M. Mitani et al., *J. Am. Chem. Soc.* **1976**, *98*, 8271–8272.
69. H. Lund, *Acta Chem. Scand.* **1963**, *17*, 972–978.
70. H. Lund, *Acta Chem. Scand.* **1963**, *17*, 2325–2340.
71. J. M. Savéant, *Acc. Chem. Res.* **1993**, *26*, 455–461.
72. B. C. Willett, W. M. Moore, A. Salajegheh et al., *J. Am. Chem. Soc.* **1979**, *101*, 1162–1167.
73. M. A. Michel, G. Mousset, J. Simonet et al., *Electrochim. Acta* **1975**, *20*, 143–149.
74. (a) B. Belleau, N. L. Weinberg, *J. Am. Chem. Soc.* **1963**, *85*, 2525–2526; (b) N. L. Weinberg, B. Belleau, *Tetrahedron* **1973**, *29*, 179–185; (c) H. Lund, J. Simonet, *C. R. Acad Sci. Paris* **1973**, *277*, 1387–1389.
75. H. Kunz, H. Waldmann, in *Comprehensive Organic Synthesis* (Ed.: B. M. Trost, I. Fleming), Pergamon, Oxford, 1991, p. 675, Vol. 6.
76. E. J. Corey, B. W. Erickson, *J. Org. Chem.* **1971**, *36*, 3553–3560.
77. B. T. Grobel, D. Seebach, *Synthesis* **1977**, 357–402.
78. P. Besse, H. Veschambre, M. Dickman et al., *J. Org. Chem.* **1994**, *59*, 8288–8291.
79. J. Gourcy, P. Martigny, J. Simonet et al., *Tetrahedron* **1981**, *37*, 1495–1502.
80. Q. N. Porter, J. H. P. Utley, *J. Chem. Soc., Chem. Commun.* **1978**, 255–256.
81. A. Lebouc, J. Simonet, J. Gelas et al., *Synthesis* **1987**, *3*, 320–321.
82. H. J. Cristau, B. Chabaud, C. Niangoran, *J. Org. Chem.* **1983**, *48*, 1527–1529.
83. P. Martigny, J. Simonet, *J. Electroanal. Chem.* **1980**, *111*, 133–137.
84. M. Platen, E. Steckhan, *Tetrahedron Lett.* **1980**, *21*, 511–514.
85. (a) L. Horner, H. Neumann, *Chem. Ber.* **1965**, *98*, 3462–3469. (b) L. Horner, R. Singer, *Justus Liebigs Ann. Chem.* **1969**, *723*, 1–10.
86. C. K. Mann, R. T. Cottrel, *J. Am. Chem. Soc.* **1971**, *93*, 3579–3583.
87. R. Kossai, G. Jeminet, J. Simonet, *Electrochim. Acta* **1977**, *22*, 1395–1402.
88. R. Kossai, J. Simonet, G. Jeminet, *Tetrahedron Lett.* **1979**, *20*, 1059–1062.
89. A. Lebouc, P. Martigny, R. Carlier et al., *Tetrahedron Lett.* **1985**, *41*, 1251–1258.
90. M. V. B. Zanoni, C. H. M. Sartorello, N. R. Stradiotto, *J. Electroanal. Chem.* **1993**, *361*, 103–108.
91. T. W. Green, P. G. M. Wuts, *Protective Groups in Organic Synthesis*, 2nd ed., Wiley, New York, 1991, pp. 46–55.
92. R. Grigg, *J. Chem. Soc., Perkin Trans* **1979**, *1*, 712 and references cited therein.
93. K. Fukase, H. Tanaka, S. Torii et al., *Tetrahedron Lett.* **1990**, *31*(3), 389–392.
94. S. Iacobucci, N. Filippova, M. d'Alarcao, *Carbohydr. Res.* **1995**, *277*, 321–325.
95. M. Platen, E. Steckhan, *Liebigs Ann. Chem.* **1984**, *9*, 1563–1576.
96. (a) P. Hodge, *Chem. Soc. Rev.* **1997**, *26*, 417–424. (b) F. Balkenhol, C. von dem Bussche-Hünnefeld, A. Lansky et al., *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2288–2337.
97. J. F. Pilard, G. Marchand, J. Simonet, *Tetrahedron* **1994**, *94*, 9401–9414.
98. (a) S. Dubey, B. Fabre, G. Marchand et al., *J. Electroanal. Chem.* **1999**, *477*, 121–129. (b) G. Marchand, J. F. Pilard, B. Fabre et al., *New J. Chem.* **1999**, *23*, 869–875.
99. W. S. Schlindwein, Y. Gofer, H. Sarker et al., *J. Electroanal. Chem.* **1999**, *460*, 46–52.

100. S. M. Hutchins, K. T. Chapman, *Tetrahedron Lett.* **1995**, *36*, 2583–2586.
101. (a) S. C. MacKeon, S. P. Watson, R. A. E. Carr et al., *Tetrahedron Lett.* **1999**, *40*, 2407–2410. (b) J. Nielsen, *Tetrahedron Lett.* **1996**, *37*, 1710.
102. (a) E. Hengge, G. Litscher, *Angew. Chem.* **1976**, *88*, 412–414. (b) T. Shono, S. Kashimura, M. Ishifune et al., *J. Chem. Soc., Chem. Commun.* **1990**, *7*, 1160–1161. (c) C. Biran, M. Bordeau, P. Pons et al., *J. Organomet. Chem.* **1990**, *382*, C17–C20.
103. E. Hengge, H. Firgo, *J. Organomet. Chem.* **1981**, *212*, 155–161.
104. C. Biran, M. Bordeau, M. P. Leger, in *Inorganic, Organometallic Polymers with Special Properties* (Ed.: R. M. Lain), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990, p. 79.
105. M. Umezawa, H. Ichikawa, T. Ishikawa et al., *Denki Kagaku* **1991**, *59*, 421–426.
106. M. Umezawa, M. Takeda, H. Ichikawa et al., *Electrochim. Acta* **1990**, *35*, 1867–1872.
107. (a) A. Kunai, T. Kawakami, E. Toyoda et al., *Organometallics* **1991**, *10*, 2001–2003. (b) P. Pons, C. Biran, M. Bordeau et al., *J. Organomet. Chem.* **1988**, *358*, 31–37.
108. M. Bordeau, C. Biran, M.-P. Leger-Lambert et al., *J. Chem. Soc., Chem. Commun.* **1991**, 1476–1477.
109. (a) F. Renauld, J. Tonge, Electrochemical synthesis of organosilicon compounds, UK Pat., GB 2 234 511A, 1991; (b) M. Bordeau, D. Deffieux, M. Leger-Lambert et al., Synthesis of organosilicon compounds, FR Pat., FR 2681866 A1, 1993.
110. M. Bordeau, C. Biran, P. Pons et al., *J. Org. Chem.* **1992**, *57*, 4705–4711.
111. V. Jouikov, G. Salaheev, *Electrochim. Acta* **1996**, *41*, 2623–2326.
112. A. Kunai, O. Ohnishi, T. Sakurai et al., *Chem. Lett.* **1995**, 1051–1052.
113. V. Jouikov, V. Krasnov, *J. Organomet. Chem.* **1995**, *498*, 213–219.
114. (a) J.-I. Yoshida, K. Muraki, H. Funahashi et al., *J. Org. Chem.* **1986**, *51*, 3996–4000. (b) V. Jouikov, L. Grigorieva, *Electrochim. Acta* **1996**, *41*, 469–470. (c) S. Torii, H. Tanaka, T. Katoh et al., *Tetrahedron Lett.* **1984**, *25*, 3207–3208.
115. A. Boukhtiarov, A. Tomilov, V. Bud'ko, *Sov. Electrochem.* **1980**, *16*, 1274.
116. M. Bordeau, C. Biran, P. Pons et al., *J. Org. Chem.* **1992**, *57*, 4705–4711.
117. D. Deffieux, D. Bonafoux, M. Bordeau et al., *Organometallics* **1996**, *15*, 2041–2046.
118. S. Grelier, T. Constantieux, D. Deffieux et al., *Organometallics* **1994**, *13*, 3711–3714.
119. S. Kashimura, Y. Murai, M. Ishifune et al., *Tetrahedron Lett.* **1995**, *36*, 4805–4808.
120. V. Jouikov, D. Fattakhova, V. Ivkov et al., *Elektrokhimica* **1996**, *32*, 297–301.
121. V. Jouikov, L. Grigorieva, *Electrochim. Acta* **1996**, *41*, 2489–2491.
122. T. Shono, S. Kashimura, H. Murase, *J Chem. Soc., Chem. Commun.* **1992**, 896–897.
123. T. Ishiwata, T. Nonaka, M. Umezawa, *Chem. Lett.* **1994**, *9*, 1631–1634.
124. C. Eaborn, R. A. Jackson, *J. Chem. Soc., Perkin Trans. 1* **1974**, *17*, 2055.
125. (a) J. Y. Becker, E. Shakkour, R. West, *Tetrahedron Lett.* **1992**, *33*, 5633–5636. (b) J. -I. Yoshida, T. Murata, S. Isoe, *Tetrahedron Lett.* **1987**, *28*, 211–214.
126. J.-I. Yoshida, T. Murata, S. Isoe, *Tetrahedron Lett.* **1986**, *27*, 3373–3376.
127. T. Takanami, K. Suda, H. Ohmori et al., *Chem. Lett.* **1987**, *7*, 1335–1336.
128. T. Koizumi, T. Fuchigami, T. Nonaka, *Electrochim. Acta* **1988**, *33*, 1635–1644.
129. S. Torii, í. Okumoto, *International Symposium on Electroorganic Synthesis*, Kurashiki, PII-29, 1994, p. 130.
130. J.-I. Yoshida, Y. Ishichi, S. Isoe, *J. Am. Chem. Soc.* **1992**, *114*, 7594–7595.
131. V. Jouikov, D. Fattakhova, *Electrochim. Acta* **1998**, *43*, 1811.
132. C. Degrand, *J. Org. Chem.* **1987**, *52*, 1421–1424.
133. S. Torii, T. Inokuchi, N. Hasegawa, *Chem. Lett.* **1980**, 639–640.
134. (a) P. Jeroschewski, W. Ruth, B. Strubing et al., *J. Prakt. Chem.* **1982**, *324*, 787–792. (b) B. Gautheron, C. Degrand, *J. Electroanal. Chem.* **1984**, *163*, 415–418.
135. S. Torii, K. Uneyama, K. Honda, *Tetrahedron Lett.* **1980**, *21*, 1863–1866.
136. A. Bewick, D. Coe, G. Fuller et al., *Tetrahedron Lett.* **1980**, *21*, 3827–3828.
137. (a) S. Torii, K. Uneyama, M. Ono, *Tetrahedron Lett.* **1980**, *21*, 2653–2654. (b) V. Jouikov, V. Latypova, Y. Kargin, *Zh. Obshch. Khim.* **1986**, *56*, 1802. (c) K. Uneyama, K. Takeno, S. Torii, *Tetrahedron Lett.* **1982**, *23*, 1161–1164. (d) K. Uneyama, H. Asai,

Y. Danoh et al., *Electrochim. Acta* **1997**, *42*, 2005–2007.
138. R. Vukicevic, S. Konstantinovic, M. Mihailovic, *Tetrahedron* **1991**, *47*, 859–865.
139. K. Uneyama, M. Ono, S. Torii, *Phosphorus Sulfur* **1983**, *16*, 35–43.
140. R. Seeber, A. Cinquantini, P. Zanello et al., *J. Electroanal. Chem.* **1978**, *88*, 137–145.
141. B. Dakova, A. Walcarius, L. Lamberts et al., *Electrochim. Acta* **1992**, *37*, 1453–1456; **1990**, *35*, 1855–1860.
142. (a) V. Jouikov, V. Ivkov, D. Fattahova, *Tetrahedron Lett.* **1993**, *34*, 6045–6048. (b) K. Surowiec, T. Fuchigami, *Tetrahedron Lett.* **1992**, *33*, 1065–1068.
143. V. Jouikov, *Zh. Obshch. Khim.* **1998**, *68*, 2071.
144. J. Simonet, J. F. Pilard, *Organic Electrochemistry*, 4th ed., Marcel Dekker, New York, 2001, pp. 1204–1208 and references cited therein.
145. K. Kyriacou, *Modern Electroorganic Chemistry*, Springer-Verlag, Berlin, 1994, pp. 215–219.
146. J. Delaunay, A. Orliac-Le Moing, A. Lebouc et al., *J. Chem. Res. (S)* **1986**, 418–419.
147. T. Masue, F. Fujihira, T. Osa, *J. Electrochem. Soc.* **1981**, *128*, 2565–2569.
148. K. Boujlel, J. Simonet, *Tetrahedron Lett.* **1985**, *26*, 3005–3006.
149. S. Nelsen, R. Akaba, *J. Am. Chem. Soc.* **1981**, *103*, 2096–2097.
150. E. L. Clennan, W. Simmons, C. W. Almgren, *J. Am. Chem. Soc.* **1981**, *103*, 2098–2099.
151. M. Schwartz, R. L. Cook, V. M. Kehoe et al., *J. Electrochem. Soc.* **1993**, *140*, 614–618.
152. G. Le Guillanton, Q. T. Do, J. Simonet, *Bull. Soc. Chim. Fr.* **1989**, *126*, 433–440; **1990**, *127*, 427–439.
153. G. Le Guillanton, Q. T. Do, J. Simonet, *Tetrahedron Lett.* **1986**, *27*, 2261–2262.
154. G. Le Guillanton, *Sulf. Rep.* **1992**, *12*, 405–435.
155. (a) Q. T. Do, D. Elothmani, J. Simonet et al., *Bull. Soc. Chim. Fr.* **1994**, *131*, 779–788. (b) D. Elothmani, Q. T. Do, J. Simonet et al., *Bull. Soc. Chim. Fr.* **1994**, *131*, 789–793.
156. Q. T. Do, J. Simonet, G. Le Guillanton, *Bull. Soc. Chim. Fr.* **1996**, *133*, 273–281.
157. Q. T. Do, D. Elothmani, J. Simonet et al., *Electrochim. Acta*, 2005, *50*, 4792–4799.
158. (a) G. Bontempelli, F. Magno, G. A. Mazzocchin, *J. Electroanal. Chem.* **1973**, *42*, 57–67. (b) A. Bewick, D. E. Coe, J. M. Mellord et al., *J. Chem. Soc., Chem. Commun.* **1980**, 51–52. (c) A. Bewick, D. E. Coe, M. Libert et al., *J. Electroanal. Chem.* **1983**, *144*, 235–250. (d) A. Bewick, D. E. Coe, J. M. Mellord et al., *J. Chem. Soc., Perkin Trans. 1* **1985**, 1033–1038.
159. S. Töteberg-Kaulen, E. Steckhan, *Tetrahedron* **1988**, *44*, 4389–4397.
160. D. Elothmani, Q. T. Do, J. Simonet et al., *J. Chem. Soc., Chem. Commun.* **1993**, 715–717.
161. N. Clauson-Kass, F. Limborg, K. Glens, *Acta Chem. Scand.* **1952**, *6*, 531–534.
162. T. Shono, Y. Matsumura, J. Hayashi et al., *Tetrahedron Lett.* **1979**, 165–168.
163. T. Chiba, H. Sakagami, M. Murata et al., *J. Org. Chem.* **1995**, *60*, 6764–6770.
164. (a) S. Torii, K. Uneyama, M. Ono et al., *Tetrahedron Lett.* **1979**, *20*, 4661–4662; (b) J. Y. Becker, D. Zemach, *J. Chem. Soc., Perkin 2* **1981**, 336; (c) K. Rossen, R. P. Volante, R. J. Reider, *Tetrahedron Lett.* **1997**, *38*, 777–778.
165. J. F. Fauvarque, G. Guérin, S. Petit et al., Process and electrolytic cell for preparing galactaric acid, Fr. Demande, Fr 2699937, 1994.
166. N. L. Weinberg, K. Hoffman, A. Kentaro, *Can. J. Chem.* **1971**, *49*, 740–745.
167. J. Gourcy, J. Simonet, M. Jaccaud, *Electrochim. Acta* **1979**, *24*, 1039–1046.
168. L. L. Miller, B. F. Watkins, *J. Am. Chem. Soc.* **1976**, *98*, 1515–1519.
169. T. Fuchigami, in *Organic Electrochemistry*, 4th ed. (Eds.: H. Lund, O. Hammerich), Marcel Dekker, NY, Basel, 2000, pp. 1035–1050.
170. J. M. Savéant, *Acc. Chem. Res.* **1980**, *13*, 323–329.
171. J. Pinson, J. M. Savéant, in *Electroorganic Synthesis. Festschrift for Manuel M. Baizer* (Eds.: R. D. Little, N. L. Weinberg), Marcel Dekker, New York, 1991, pp. 29–44.
172. J. Delaunay, A. Orliac, J. Simonet, *Tetrahedron Lett.* **1986**, *27*, 6205–6206; **1995**, *36*, 2083–2084.
173. J. Delaunay, J. Simonet, *C. R. Acad. Sci. Paris* **1997**, *324II*, 269–273.
174. M. F. Nielsen, B. Batanero, T. Lohl et al., *Chem. Eur.* **1997**, *3*, 2011.
175. D. Caras-Quintero, P. Baeuerle, *Chem. Commun.* **2004**, *8*, 926–927.
176. A. M. Marthe, G. Mousset, P. Pouillen et al., *Electrochim. Acta* **1991**, *36*, 1911.

177. L. Horner, D. Degner, *Tetrahedron Lett.* **1968**, 5889.
178. A. R. Amundsen, E. N. Balko, *J. Appl. Electrochem.* **1992**, *22*, 810.
179. H. Simon, H. Gunther, J. Bader et al., *Angew. Chem.* **1981**, *93*, 897.
180. R. Yuan, S. Watanabe, S. Kuwabata et al., *J. Org. Chem.* **1997**, *62*, 2494–2499.
181. E. Dunach, A. P. Esteves, M. Medeiros et al., *J. Electroanal. Chem.* **2004**, *566*, 39–45.
182. J. Pytkowicz, S. Roland, P. Mangeney et al., *J. Organomet. Chem.* **2003**, *678*, 166–179.
183. M. Durandetti, J. Perichon, J.-Y. Nedelec, *J. Org. Chem.* **1997**, *62*, 7914–7915.
184. S. Quideau, I. Fabre, D. Deffieu, *Org. Lett.* **2004**, *6*, 4571–4573.
185. H. Putter, Industrial electroorganic chemistry, *Organic Electrochemistry*, 4th ed., Marcel Dekker, New York, 2001.
186. I. M. Dalrymple, A. R. Jones, J. P. Millington, *Electrochemical Society Meeting*, Extented Abstract, Washington, DC, 1976, p. 690.
187. E. Steckhan, G. Hilt, R. Kempf et al., Organic synthesis, Proc. Symp., 5th, Heidelberg, Sept. 26–28, 1996.
188. D. Pletcher, F. Walsh, *Industrial Electrochemistry*, 2nd ed., Chapman & Hall, London, 1990, pp. 313–323.
189. M. M. Baizer, *J. Electrochem. Soc.* **1964**, *111*, 215.
190. M. M. Baizer, D. E. Danly, *CHEMTECH* **1980**, *10*, 302.
191. D. Degner, H. Hannebaum, H. Nohe et al., *AlChE Symp Ser.* **1979**, *75*, 14–18.
192. M. Seko, A. Yomiyama, T. Isoya, *Hydrocarbon Process.* **1979**, 117.
193. H. Putter, J. Botzem, H. Hannebaum et al., Elektrochemische Methoxylierung von Ketronen, *Annual Conference of the GDCh. Fachgruppe Angewandte Elektrochemie*, Frankfurt, September 1998, GDCh Monograph, pp. 458–465.
194. BASF information 15.10.98, Company brochure, BASF, Ludwigshafen.
195. P. Vaudano, E. Plattner, C. Comninellis, *Chimia* **1995**, *49*, 12–16.
196. (a) S. Dapperheld, R. Rossmeissel, Process for partial electrolyte dehalogenation of dichloro- and trichloroacetic acids and electrolysis solutions, Geo. Offen. DE 4029068, 1992.
197. J. Chaussard, M. Troupel, Y. Robin et al., *J. Appl. Electrochem.* **1984**, *19*, 345–348.

7
Batteries: Basic Principles, Technologies, and Modeling

Gerardine G. Botte
Ohio University, Athens, OH

7.1	**Introduction**	**379**
7.2	**Basic Principles**	**379**
7.3	**History of Batteries**	**384**
7.4	**Battery Applications and Market**	**388**
7.5	**Thermodynamics of Batteries and Electrode Kinetics**	**391**
7.5.1	Thermodynamics and Cell Potentials	391
7.5.2	Electrode Kinetics	396
7.6	**Transport Mechanisms in Batteries**	**397**
7.7	**Characteristics of Batteries**	**398**
7.8	**Theoretical Capacity and Voltage**	**400**
7.8.1	Theoretical Capacity	400
7.8.2	Theoretical Voltage	402
7.9	**Battery Technologies**	**402**
7.9.1	Primary Batteries	402
7.9.1.1	Leclanché's Cells	402
7.9.1.2	Magnesium Cells	404
7.9.1.3	Alkaline Manganese Dioxide Batteries	405
7.9.1.4	Silver Oxide Cells	406
7.9.1.5	Zinc/Air Cells	406
7.9.1.6	Lithium Batteries	407
7.9.1.6.1	Li/SO_2 Cells	407
7.9.1.6.2	Li/MnO_2 Cells	408

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

7.9.2 Secondary Batteries 409
7.9.2.1 Lead-acid Batteries 409
7.9.2.2 Nickel–Cadmium Batteries 411
7.9.2.3 Nickel–Hydrogen Batteries 412
7.9.2.4 Nickel–Metal Hydride (Ni/MH) Batteries 413
7.9.2.5 Lithium-ion Batteries (Li-ion) 414

7.10 Mathematical Modeling of Batteries 415
7.10.1 Schematic Diagram and Complexity of the Model 416
7.10.1.1 Empirical Models 416
7.10.1.2 First-principle Models 417
7.10.2 Formulation of the Equations 417
7.10.3 Solution of Model Equations 418
7.10.4 Estimation of Properties and Parameters 418
7.10.5 Validation of the Model 418

7.11 Summary and Battery Trends 418

7.12 List of Symbols 420
Greek 420
Subscripts 421
Superscripts 421
References 421

7.1 Introduction

Many things occur to us when we hear the word battery. The thoughts about 'battery' depend on the person: a child will think about a favorite toy that moves or talks or does both, a teenager will think about a CD player or game, an executive will remember his cell phone or possibly his laptop, a mechanic will see a car, and a medical doctor will see a pace maker or a hearing-aid device. The different thoughts are examples of the applications of batteries across the world. Batteries are widespread all over the world; they can be as small as the coin batteries used to provide power for watches and hearing-aid devices, or as large as the ones used to provide power for satellites and electric vehicles. Independent of the size, a battery is a complex device that is able to transform chemical energy into electrical energy. It is called *complex* because it involves many areas of science, such as chemistry, kinetics, transport phenomena, thermodynamics, and physics. Consequently, the development of battery systems is a very challenging process.

In this chapter, an introduction to batteries is presented including an understanding on how they work; their applications; their thermodynamics; the kinetics of their electrodes; the transport processes that occur in batteries; and the fundamental variables involved in their design. An overview of the different battery technologies available in the market, an explanation of how to predict or model battery behavior, and some applications of battery modeling are also included. Finally, the future directions in battery technologies are discussed.

7.2 Basic Principles

This section presents the definition of specific terms used in battery technology and the basic principles for the regular operation of batteries. Lithium-ion batteries will be used as an example to illustrate these concepts because they possess attributes that exhibit the most recent technologies available in batteries.

A battery is a complex device that delivers electrical energy by transforming chemical energy. The electrical energy is provided by electrochemical reactions (oxidation–reduction reactions) that take place at the anode and the cathode of the battery. While the term battery is often used, the basic electrochemical unit being referred to is the cell [1]. A battery is composed of several cell units that are connected in series or in parallel

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

depending on the desired voltage and capacity. Series connections are used to provide a higher voltage than the voltage of one cell, while parallel connections are used to provide a higher capacity than the capacity of one cell.

Batteries can be classified according to their application, as primary and secondary batteries. Primary batteries are not rechargeable while secondary batteries are rechargeable. Figure 1 shows a schematic of a unit cell for a lithium-ion battery, which is a secondary type of battery. The figure shows the main components of a unit cell: negative electrode, positive electrode, separator, electrolyte, binder, filler, current collectors, and tabs. All these components are typical parts of any battery; however, the specific materials used for all of them depend on the battery type and application. A description of each of the components is given next.

Figure 1 shows the negative and the positive electrodes in a lithium-ion cell. The

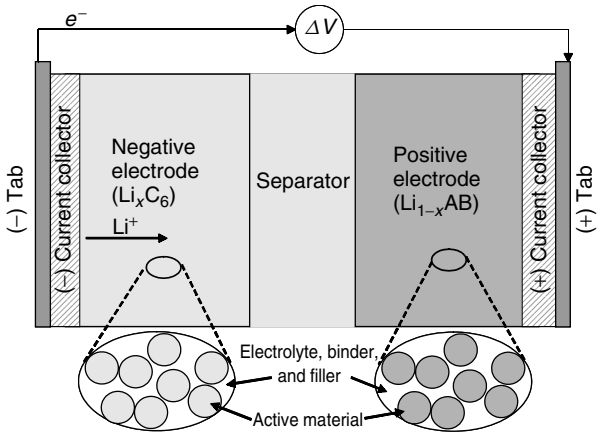


Fig. 1 Idealized schematic diagram of a lithium-ion cell during discharge. The cell sandwich consists of a negative lithiated carbon electrode Li_xC_6 (e.g. natural graphite, petroleum coke), separator (e.g. polyethylene, polypropylene), and a positive electrode $Li_{1-x}AB$ ($Li_{1-x}AB$ represents a lithiated composite material, e.g. $Li_{1-x}CoO_2$, $Li_{1-x}NiO_2$, $Li_{1-x}MnO_2$). The amount of lithium present in the electrodes in the intercalated state is represented by x, and $1 - x$, for the negative and positive electrodes, respectively. The active material in the electrodes is depicted as spherical particles (active material shapes vary, e.g. spherical, fibers, amorphous). During the discharge, lithium is deintercalated from the negative electrode and oxidized at the surface of the particle-producing lithium ions (see Eq. 1). Lithium ions migrate and diffuse to the positive electrode where they are reduced at the surface of the particle-producing lithium in the intercalated state (see Eq. 2). The electrons (generated from the cell reaction at the negative electrode, Eq. 1) flow from the negative electrode to the positive electrode (where they participate in the cell reduction process, Eq. 2) through the external circuit causing a difference in the potential (ΔV). During the charge of the cell, the opposite processes take place.

negative electrode is a type of lithiated carbon, petroleum coke, or graphite (Li_xC_6), while the positive electrode is a lithium composite material that has been represented by $Li_{1-x}AB$ (e.g. $Li_{1-x}CoO_2$, $Li_{1-x}NiO_2$, $Li_{1-x}MnO_2$). In the battery nomenclature, the electrode that is defined as the anode during the discharge of the battery (or cell) is called the *negative electrode* (it has a negative charge because electrons are being produced), while the electrode that is defined as the cathode during the discharge of the battery is called the *positive electrode* (it has a positive charge because electrons are being consumed). It is important to mention that the anode is not always called the negative electrode, and the cathode is not always called the positive electrode; this is true in primary batteries, and in secondary batteries only during the discharge process. In secondary batteries, the designation of the anode and the cathode electrodes depends on the discharge or charge process but the negative and positive electrodes are always the same by definition. The nomenclature of negative and positive electrodes was defined to avoid confusion of the electrodes in secondary batteries. For this reason, it is preferable to use the terms negative and positive electrodes in secondary batteries, instead of anode and cathode. At the anode (also known as the *reducing electrode*) an oxidation process takes place (loss of electrons); in the case shown in Fig. 1, the lithium intercalated into the carbon is being oxidized into lithium ions and electrons are produced according to the oxidation reaction [2]:

$$Li_xC_6 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} xLi^+ + C_6 + xe^- \quad (1)$$

where x represents the lithium fraction deintercalated during the process. At the cathode (also known as the *oxidation electrode*) a reduction process takes place; in the case shown in Fig. 1, lithium ions are reduced into lithium in the solid-intercalated state in the specific material. Electrons are being consumed in this process according to the reduction reaction:

$$xLi^+ + xe^- + Li_{1-x}AB \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} LiAB \quad (2)$$

where AB represents a composite material. The reactions given in Eqs (1) and (2) are reversible, that is the reason why lithium-ion batteries are rechargeable. Theoretically, the deintercalation and intercalation fractions of lithium (x) in both the electrodes should be the same (as shown in Eqs 1 and 2). However, because of side reactions or decomposition reactions that take place during the charge and discharge processes lithium may react with other species present in the battery affecting the lithium balance shown in Eqs (1) and (2) (e.g. the formation of the solid electrolyte interface, SEI). These details have been suppressed to avoid complications that are not necessary to illustrate the objectives of this section. More details about side reactions in lithium-ion batteries can be found in the literature [3].

There are two types of electrodes, metal electrodes and porous electrodes. For instance, lithium-ion batteries possess porous electrodes (as shown in Fig. 1), while lead-acid batteries have metal electrodes. The selection between metal and porous electrodes depends on the chemistry and on the safety of the battery. Porous electrodes in a battery are usually very thin to reduce the internal resistance of the battery. They are typically made of small particles to reduce internal diffusion limitations and to provide a large

surface area for electrochemical reactions. In Fig. 1 the active material in the electrodes is depicted as spherical particles; however, it can have different shapes (e.g. fibrous, amorphous). Independent of the electrode type, the capacity of the positive electrode is normally different from that of the negative electrode in a battery. If there is extra capacity on the positive electrode, the battery is negative limited, and vice versa. For example, nickel electrode–based batteries and lithium-ion batteries are normally positive limited.

The oxidation and reduction reactions (e.g. Eqs 1 and 2, respectively) take place at two different sides of the battery or cell, allowing the electrons to flow through an external circuit that is connected to the specific application. For the battery to operate, the ions need to be transported inside the cell. The transport of ions is done through the electrolyte. The electrolyte is another component of the unit cell as shown in Fig. 1. The ions in the electrolyte can move through diffusion, migration, and convection (although convection is usually negligible in batteries because the velocity of the ions is very slow). A good electrolyte needs to have a high ionic conductivity (to allow fast transport of the ions reducing the resistances inside the cell or battery), it needs to be chemically stable with the electrodes and other components, and it should also be thermally stable at the operating temperature of the cell or battery. The electrolytes can be liquids or solids, and aqueous or nonaqueous. The selection of the electrolyte system depends on the chemistry of the battery.

The anode and the cathode cannot be in direct contact; otherwise the battery would be short circuited. To avoid direct contact of the electrodes a separator is used between them, as shown in Fig. 1. The separator should allow the transport of the ions very easily, therefore the material should be permeable to the ions and it should also have a high ionic conductivity. The separator should be chemically stable with the electrolyte, the active materials, and other components at the temperature of operation of the cell or battery.

Other components of the battery are the current collectors. Figure 1 shows the current collectors for the positive and negative electrodes in the lithium-ion cell. The active materials for the positive and negative electrodes (in this case) consist of thick layers of porous materials coated onto the current collectors. The current collector works as a support for the active material and provides a conducting path for the active material paste reducing the resistance of the battery [4]. Current collectors are used only in porous electrodes; metal electrodes do not require current collectors. The current collectors should be chemically stable and resistant to corrosion. They should also have a high electronic conductivity to reduce the internal resistance of the battery. Copper and aluminum are the current collectors used in lithium-ion batteries for the negative and positive electrodes, respectively.

Attached to the current collectors of the battery (or to the electrode in metal electrodes) are the tabs. The tabs electrically connect the battery to the can, to transport the electrons and therefore the current. They should also be chemically stable and resistant to corrosion. At the same time, they should have a high electronic conductivity to reduce the internal resistance of the battery. The tabs used in lithium-ion batteries are made of nickel. Other additional components shown in Fig. 1 are the binder and the filler. The binder is the chemical compound used in this case to attach the active material to the current

collectors. Polyvinylidene fluoride (PVDF) is the compound used as binder in lithium-ion batteries. The filler is a compound used to increase the electronic conductivity of the electrodes and decrease the internal resistance of the cell. Carbon black is the most common filler used in lithium-ion batteries. Both binder and filler are used in porous electrodes only.

Figure 2 shows a schematic diagram of a complete battery (lithium-ion battery). Several cells as the one shown in Fig. 1 are wrapped together in parallel. The schematic corresponds to a cylindrical battery. The anode, separator, and cathode materials are tightly wrapped and held together to form what is called the *jelly roll* (see Fig. 2(a)). The jelly roll is introduced into the container or can. The container or can should be resistant to corrosion from both inside and outside. It should also have the required mechanical strength for the specific application [4]. The containers usually have a plastic insulator for protection of the can from the external media.

More details of the components of the battery including safety devices are given in Fig. 2(b). Electrical energy is stored in a battery in the form of chemical energy; this makes the battery susceptible to safety problems that need to be controlled through safety devices. Three safety devices are shown in Fig. 2(b), the burst disk, the positive temperature coefficient (PTC), and the separator. One of the most common safety devices is the burst disk. The burst disk is an irreversible safety mechanism. When the pressure of the battery increases under extreme conditions, the disk breaks releasing the pressure of the battery. The PTC is a reversible safety mechanism. Its function is to stop the current flow if the battery overheats. When the battery overheats, the material inside the PTC expands disconnecting the header of the battery from the positive tab [5]. The process is reversible, and after the battery cools down, the material inside the PTC is compressed, electrically connecting the battery. The separator also acts as a safety device. When the temperature of the battery undergoes extreme conditions, the separator melts thereby stopping the flow of ions through the battery, and

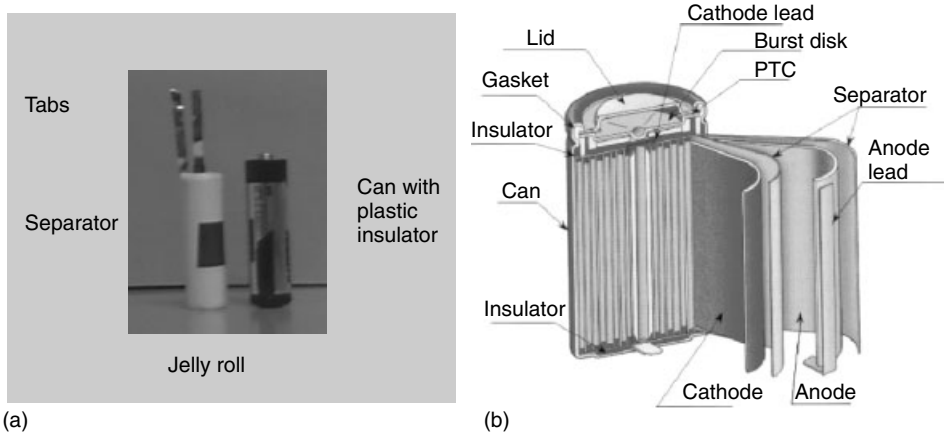


Fig. 2 Schematic diagram of a complete lithium-ion battery. (a) Jelly roll, (b) other components and safety devices (from Ref. 5).

therefore the flow of current. This process is irreversible. Similar safety devices can be found in other types of batteries.

Batteries are made in three special shapes: cylindrical (as shown in Fig. 2), flat or prismatic, and button shape. The selection of the shape depends on the application, for example button cells are used to power hearing-aid devices. The three special shapes are available in different sizes. Batteries are produced by different manufactures all around the world. Independent of the manufacturer, batteries should meet international standards that make them interchangeable. The international criteria that standardize batteries have been prepared by the International Electrotechnical Commission (IEC). The standardization of batteries includes the following criteria: dimensions (sizes, shape), voltage, polarity, terminals, nomenclature, and marking. The standardization of batteries not only make the batteries from different manufacturers interchangeable but also (1) limits the number of battery types, (2) defines the standard quality of batteries and provides guidance for its assessment, and (3) supplies guidance on safety matters during transportation, storage, and uses of batteries. Battery standardization allows a third-party buyer to easily replace a battery for a specific application. More details about battery standardization are given elsewhere [1].

7.3 History of Batteries

Batteries were probably used more than 4000 years ago. It is believed that Parthians used batteries to electroplate silver (Baghdad, 250 B.C.), and also that the Egyptians were able to electroplate antimony onto copper more than 4000 years ago [6]. However, the written evidence in battery history began 200 years ago. More than two centuries have passed since Luigi Galvani performed the first experiment related to batteries (frog leg experiment, 1791). His experiments were continued by Alessandro Volta (voltaic pile, 1800) [7]. Since then, a lot of changes and new technologies in batteries have been developed and introduced into the market. Figure 3 presents a schematic of the major events during the evolution of batteries in time, starting with Galvani's frog leg experiment (1791) and finishing with the most recent technologies (lithium-ion batteries, 1991).

Luigi Galvani (1791) was the first to discover the physiological action of electricity. He demonstrated the existence of bioelectric forces in animal tissue. His experiments led Alessandro Volta to the invention of the first battery, voltaic pile [8]. In 1800, Alessandro Volta described the voltaic pile in a letter to the Royal Society in London [7]. The original voltaic cell used two metal disks as electrodes, namely zinc and silver. Cardboard disks separated the electrodes and seawater was the electrolyte. A current was produced when the silver disk was connected to the zinc disk through an external wire. The voltaic pile established the foundation for the liquid battery type. Many other metals and electrolytes have been tried during the last two centuries [9].

John Frederic Daniell started some experiments to improve the behavior of the voltaic pile. The use of the voltaic pile was limited because it was not able to sustain a constant current for a certain period of time. The problem was due to the polarization of the cell. In 1836, Daniell improved the voltaic cell by solving the polarization problem. He invented a cell in which hydrogen was eliminated through the generation

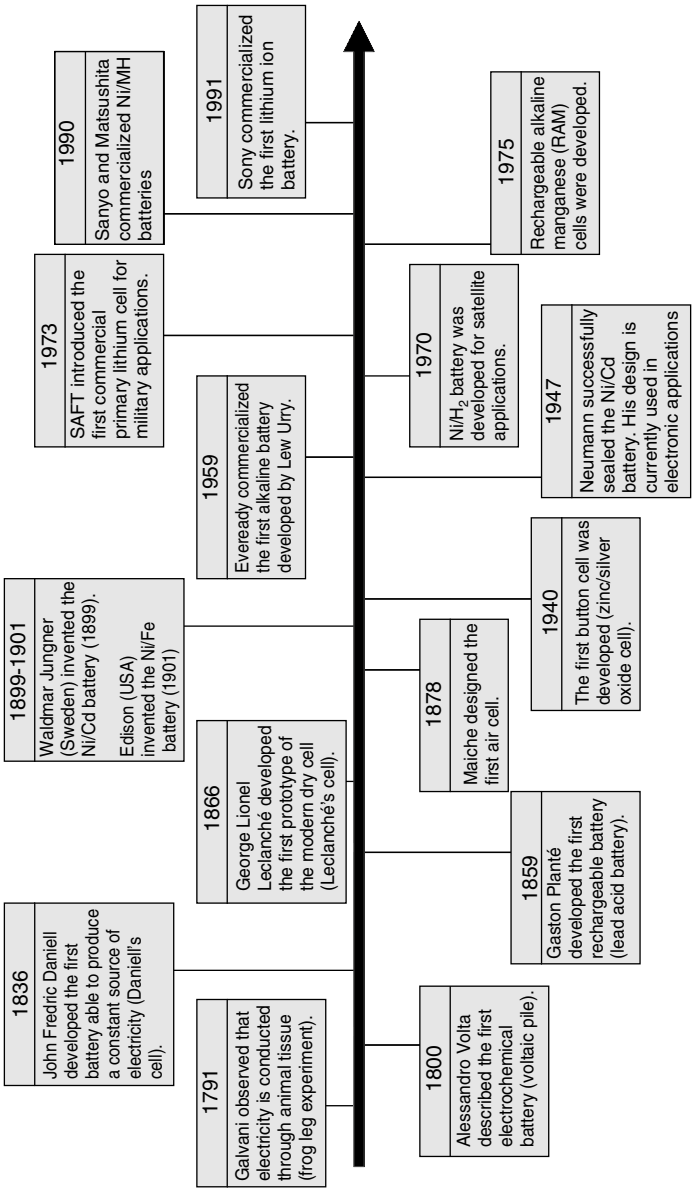


Fig. 3 Timeline for the major events in the history of batteries.

of electricity. Daniell is responsible for inventing the first battery able to produce a constant source of electricity for a long period of time (Daniell's cell [9]). Daniell's cell used copper and zinc as the working electrodes, porous earthenware as separator, and copper sulfate as electrolyte.

Research in batteries kept going on for specific applications. In 1859, Raymond Gaston Planté developed the first rechargeable battery, lead-acid battery (also known as *Planté's cell* in honor of its inventor). His design was developed further by Faure in 1881 [10]. The original Planté cell consisted of two sheets of soft lead plates, separated by rubber strips, rolled into spirals, and immersed in a solution of 10% sulfuric acid [8]. This cell led to the conversion of lead oxide (PbO_2) at the surface of the positive electrode, and metallic spongy lead at the negative electrode. The capacity of the Planté battery was limited because the positive electrode had a small amount of active material. In 1881, Faure modified Planté's cell. He used a paste of lead oxide or pasted plate (a sheet of lead coated with lead oxide and sulfuric acid) for the positive electrode instead of the soft lead plate used by Planté. This change in the electrode provided more mass of active material, increasing the capacity and the efficiency of the cell. Even though this cell was invented more than 150 years ago, it is still one of the most important secondary batteries in the market.

Georges-Lionel Leclanché developed the first prototype of the modern dry cell, Leclanché's cell, in 1866. The cell was designed with the purpose of providing a reliable and easily maintainable power source for telegraphic offices [1]. The original cell was assembled in a porous pot that worked as the separator [8]. The positive electrode consisted of manganese dioxide with a little mix of carbon, the negative electrode was a zinc rod, and the electrolyte was a solution of ammonium chloride. Leclanché's cell became the first prototype of the dry zinc–carbon cell. This cell was inexpensive, safe, easy to maintain, and provided an excellent storage life with the appropriate performance [1]. In 1876, Leclanché improved the design of his cell by removing the porous pot and by adding a resin to the manganese dioxide [1]. Other researchers such as Carl Gassner (1888) worked on a better prototype of Leclanché's cell, resulting in the construction of dry cells.

Other changes were also made in Leclanché's prototype. In 1878, Maiche designed the first air cell by replacing the manganese dioxide cathode of Leclanché's cell with a mixture of platinum carbon powder/air electrode [7]. Limitations in his technology delayed the commercialization of the zinc/air cells. It was not until 1932, when Heise and Schumacher modified the cell design, that the air cells were commercialized [1]. They built alkaline electrolyte zinc/air cells, which used porous carbon air cathodes impregnated with wax to prevent flooding. Their design has practically remained unchanged and is still in use for the manufacture of large industrial zinc/air batteries. These batteries use oxygen (cathode reactant) from the atmosphere to produce electrochemical energy.

In 1899, Waldmar Jungner (Sweden) invented the nickel–cadmium battery (Ni/Cd) [6]. The battery used nickel for the cathode, cadmium for the anode, and an aqueous solution of potassium hydroxide for the electrolyte. The applications of these batteries were limited because of the high cost of the materials compared to other battery systems (lead acid). In 1901, Edison (USA) modified the design of the battery by replacing the anode material with iron [7]. This design is known as the

nickel iron battery (Ni/Fe). The applications of these batteries were also limited because of the cost of the materials. Many improvements have been made in these batteries, particularly associated with the method of production of the nickel oxide electrodes. In 1912, Edison patented a sealed Ni/Fe cell but the commercialization of sealed cells started 50 years later in Europe [7]. In 1932, Slalecht and Ackermann (Germany) invented the sintered pole plate [6]. The electrodes produced according to this method had a low internal resistance producing large improvements in the performance of the batteries. It was not until 1947 that the sealed Ni/Cd battery [6], which is the design that we currently use in applications such as computers, telecommunications, electronics, and so forth, was designed by Neumann (France).

More research was done on electronic applications, resulting in the development of the button cells and the alkaline batteries. The introduction of the button cells represented a significant advance for the electronic business. The first button cell to be developed was the zinc/mercuric oxide cell, in the 1940s. Later the zinc/silver oxide cell was developed, in the 1960s [10]. Both cells used amalgamated pellet zinc powder as anode, HgO (zinc/mercury cell) or Ag_2O_2 (zinc/silver oxide cell) mixed with carbon as cathode materials [10], and an alkaline solution (potassium hydroxide in water) as electrolyte. Another significant advance for the electronic business was the commercialization of the alkaline batteries. In 1959, Eveready commercialized the first alkaline battery. The battery was developed by Lew Urry [8]. The battery is a variation of Leclanché's cell. It uses manganese oxide (MnO_2) for the cathode material, powdered zinc for the anode material, and potassium hydroxide with some quantities of zinc oxide in aqueous solution for the electrolyte [7]. These batteries are now available in two different sizes: relatively large cylindrical cells and miniature button cells.

Long cycle batteries were needed for satellite applications. In 1970, the first nickel–hydrogen battery (Ni/H_2) was developed for this purpose. These batteries are able to store hydrogen at very high pressures but they are very expensive. The long cycle capability of these batteries catalyzed the research and development of a similar type of battery for consumer applications. Different alloys were developed to store hydrogen; however, they were very unstable. It was not until the 1980s, when enough stable materials were developed for use in commercial cells [6], that the nickel–metal hydride cells (Ni/MH) were designed. In 1990, a Ni/MH battery was commercialized in Japan by Sanyo and Matsushita [11]. These two companies started the large-scale commercialization of small sealed Ni/MH batteries. The companies associated themselves with Duracell, Toshiba, and Varta in a consortium known as the *3C alliance* (camcorders, cellular phones, and computers) [7]. The idea of the ''3C alliance'' was to use the Ni/MH batteries for applications such as communications, cordless consumer products, and other high rate long cycle life applications. This cell system is currently under research and development for electric vehicles applications.

A revolutionary event in the primary battery market was the development of primary lithium batteries in 1973. Lithium is a very attractive metal for battery applications due to its low atomic mass (6.94), its high specific capacity (3.86 Ah g^{-1}), and its high electrochemical reduction potential (−3.035 V) [10]. Research in lithium batteries started in the late 1950s. It was then

demonstrated that lithium metal could be stable in a series of electrolyte systems due to the passivation of the material [7]. Lithium reacts with the electrolyte system forming a passivation layer called *solid electrolyte interface* that makes the material chemically stable. In 1973, SAFT introduced the first commercial lithium cell for military applications. Its use was limited because of some safety considerations [1]. Today, primary lithium batteries are largely used in electronic applications.

Continuous research in primary alkaline manganese batteries ended up in the development of rechargeable alkaline manganese (RAM) cells. The design of these cells dated to 1975 [1]. These batteries are basically an extension of the primary alkaline batteries. They also use zinc for the negative electrode, manganese dioxide for the positive electrode, and an aqueous solution of potassium hydroxide for the electrolyte.

One of the major events in the history of modern batteries was the development of secondary lithium batteries. Attempts to develop these types of batteries started during the 1980s [6]. However, the attempts failed because of the safety problems of the battery. The first rechargeable lithium batteries used lithium foil as the negative electrode, a lithiated metallic oxide intercalation compound as the positive electrode, and a liquid organic electrolyte. The applications of these batteries were limited to small scale size because of safety concerns; under certain conditions the temperature of the battery increased violently, causing fire and explosion of the battery. This phenomenon is known as *thermal runaway*. It was found that during cycling of the battery, lithium was deposited on the negative electrode (on the lithium metal) forming a rough shape on the surface of the material. The lithium metal electrode formed a rough surface susceptible to violent reactions. In 1991, Sony commercialized the first lithium-ion battery, which uses a carbon intercalation material for the negative electrode instead of lithium metal. This design tremendously improved the safety of these batteries by reducing the lithium metal deposition. Today, the lithium-ion batteries have invaded the electronic applications market and they are under research and development for large-scale applications such as electric vehicles and satellites.

7.4 Battery Applications and Market

Batteries are used in many different fields: consumer, electronics, telecommunications, military and defense, space, and medicine. They are designed according to the specific application, and can be classified, as mentioned in Sect. 7.2, into primary and secondary batteries. The applications of both primary and secondary batteries are enormous; some examples are given in Table 1.

It is known that the battery market has been in a continuous growth as evidenced by the increase in the number of portable applications that use batteries, especially electronic applications (see Table 1). However, there is no definitive authority reporting on world battery markets, and the majority of the marketing analysis in batteries are private reports specializing in a particular area. Therefore, the analysis presented in this section is only based on the reputed information available in the unrestricted literature, which is basically based on battery sales until 1997 (information for further years is not available in the public literature). Since Japan is one of the countries with the largest battery production in

Tab. 1 Examples of battery applications

Battery	*Applications*
Primary batteries	*Portable electronic devices*: watches, cameras, camcorders, calculators, remote controls, walkmans, toothbrushes, shavers, toys. *Household*: flash lamps, smoke detectors, radios. *Workshop*: portable tools (wireless drill, screwdrivers, etc.), test meters. *Medicine*: hearing-aid devices, pace maker devices, blood pressure monitors.
Secondary batteries	*Portable electronic devices*: cell phones, cordless phones, laptops, camcorders, toys. *Communications*: satellites, military radios. *Transportation*: electric vehicles, car ignition batteries. *Defense*: torpedoes, missiles.

the world, information about their market is used as reference in this analysis. Figure 4 shows the sales of small-size batteries in Japan from 1992 to 1997 [12]. The sales followed a parabolic growth tendency. In 1997, the total worldwide sales of batteries were about $14 billion [13], while the total battery sales in Japan were $6.74 billion, which represent 48% of the global battery sales in the world. The percentage left (52%) was distributed among US, China, Indonesia, Singapore, UK, and other countries in Europe.

In 1997, 25% of the battery sales in Japan corresponded to primary batteries ($1.69 billion) and the other 75% corresponded to secondary batteries ($5.05 billion) [12]. It is expected that in the coming years the sales in secondary batteries will increase even more and the sales in primary batteries will decrease, as secondary batteries become less expensive and will substitute some of the primary batteries applications. Figure 5 shows a breakdown of the battery sales in Japan during 1997. The major sales in primary batteries corresponded to alkaline manganese batteries (40% of $1.69 billion). The major sales in secondary batteries corresponded to lithium-ion batteries (34% of $5.05 billion), even though the production of these batteries was smaller than the production of Ni–Cd and Ni–MH batteries for that year, as evidenced from Fig. 6. The difference was basically due to the higher cost of lithium-ion batteries compared to that of Ni–Cd and Ni–MH batteries. More than half of the sales (59%) in lead-acid batteries were mainly for automobiles excluding motorcycles, and less than half of the sales were for industries (41%).

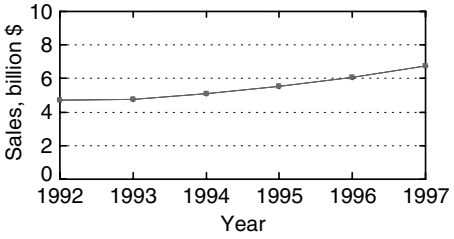


Fig. 4 Battery sales in Japan from 1992 to 1997 (from Ref. 12). The sales showed a parabolic trend. In all the years, the sales showed a positive marked growth with respect to the previous year.

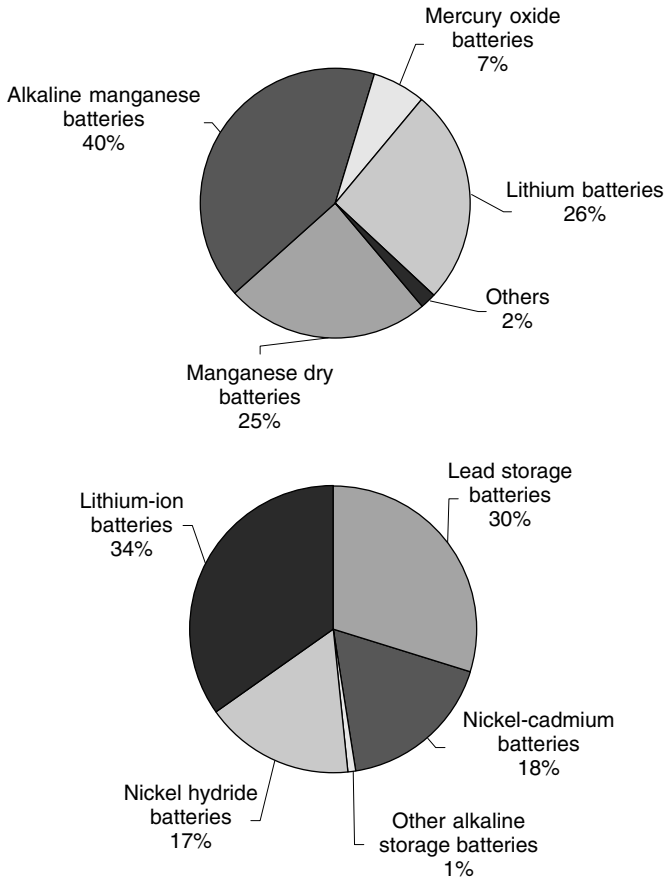


Fig. 5 Battery sales breakdown in Japan during 1997; from a total of $6.74 billion, $1.69 billion (25%) corresponded to primary batteries and $5.05 billion (75%) corresponded to secondary batteries (from Ref. 12).

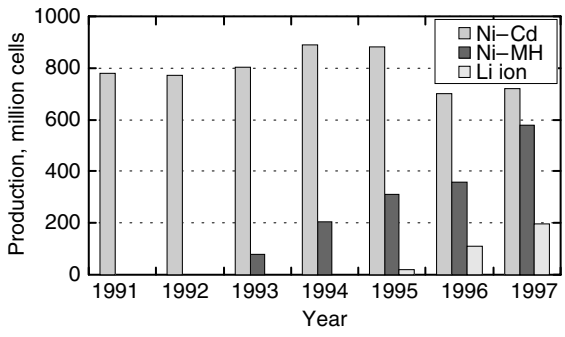


Fig. 6 Production of small-size secondary batteries in Japan from 1991 to 1997 (from Ref. 11).

Figure 6 presents the production of small-size secondary batteries in Japan from 1991 to 1997 [11]. The production of Ni–Cd batteries started to decrease with the introduction of the Ni–MH and Li-ion batteries. It is expected that the production of lithium-ion batteries will increase in the coming years taking over the majority of the market. Nowadays, lithium-ion batteries are the preferred ones for laptops, while Ni–MH batteries dominate the market for cordless phones.

7.5 Thermodynamics of Batteries and Electrode Kinetics

The thermodynamics of batteries and the kinetics of their electrodes are very complex. The purpose of this section is to introduce the reader to the basic concepts in this area in a simple manner. To achieve this goal, the concepts will be introduced using a simple electrochemical cell as an example, which uses potassium and mercuric oxide as electrodes, instead of using the lithium-ion battery described in Sect. 7.2, which is a more complex system. It is impossible to capture all the details in such a short account. A more comprehensive analysis is available elsewhere [14, 15].

7.5.1 Thermodynamics and Cell Potentials

Suppose that we want to build an electrochemical cell (a single unit in a battery) using a potassium electrode, a mercuric oxide electrode, and potassium hydroxide in water as electrolyte. The oxidation and reduction reactions are

$$\mathrm{K(s)} \longrightarrow \mathrm{K^+} + e^- \tag{3}$$

$$\mathrm{HgO} + \mathrm{H_2O} + 2e^- \longrightarrow \mathrm{Hg(l)} + 2\mathrm{OH^-} \tag{4}$$

The cell will be connected to the external circuit using platinum tabs. The schematic of the cell is given in Fig. 7. Several questions about the cell can be formulated: (1) What is the polarity of the electrodes? (2) What is the maximum potential that the cell can deliver? (3) Does the concentration affect the potential of the cell? One way to address these questions is experimentally, by connecting the electrodes to a voltmeter.

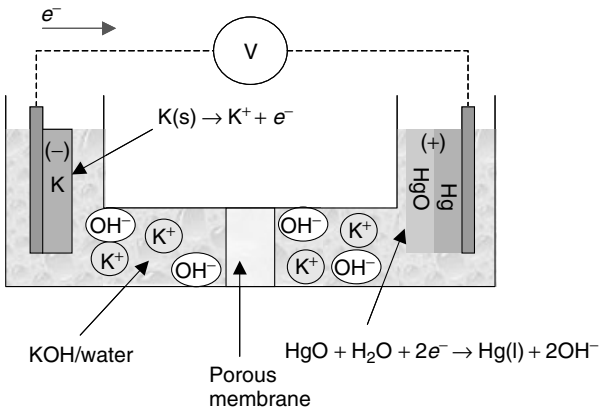


Fig. 7 Schematic diagram of a K/HgO–Hg cell.

The voltmeter will indicate the maximum voltage that the cell can deliver. To deliver energy from the cell (which is the function of a battery), the polarity of the voltage should be positive; this criterion is used to connect the electrodes properly to the external circuit (to decide which are the positive and the negative electrodes). Consequently, the connections of the electrode tabs should be switched until a voltage with a positive polarity is measured. In our example, the negative electrode is the potassium electrode and the positive one is the mercuric oxide electrode, and this connection to the voltmeter should give a positive potential. The potential measured in the voltmeter as soon as the cell is connected to it should be 3.013 V. This potential is known as the *standard open-circuit potential*. It is called *open* because there is no current passing through the external circuit (realistically when the system is closed through the voltmeter, there will be a small current very close to zero). The potential measured will be the maximum potential of the cell and it is independent of the concentration. If the circuit is kept open for a long time at a constant temperature the potential will drop because of the change in the concentration of the cell. All these issues can also be addressed theoretically using thermodynamics.

Thermodynamics can be used to predict the open-circuit potential of the cell, and the reading of the voltmeter (standard potential) when only reversible effects are considered (equilibrium conditions) by using thermodynamic equilibrium relations. Nonreversible effects (i.e. diffusion and migration) can be predicted using transport phenomena. The first step to apply the thermodynamic equilibrium relations is to draw a schematic diagram of the cell distinguishing the different phases as given by

$$\left|\begin{array}{c|c|c|c|c|c|c} \alpha & & \beta & & \delta & & \\ \mathrm{Pt} & & \mathrm{K(s)} & & \mathrm{KOH/H_2O} & & \\ & \varepsilon & & \phi & & \alpha' & \\ & \mathrm{HgO} & & \mathrm{Hg(l)} & & \mathrm{Pt} & \end{array}\right| \tag{5}$$

where lines separate the different phases. When the cell is in equilibrium, the following should be true

- Phase α in equilibrium with phase β
- Phase β in equilibrium with phase δ
- Phase δ in equilibrium with phase ε
- Phase ε in equilibrium with phase ϕ
- Phase ϕ in equilibrium with phase α'.

The equilibrium condition at each phase is represented by

$$\mu_i^{\mathrm{Phase1}} = \mu_i^{\mathrm{Phase2}} \tag{6}$$

where μ represents the chemical potential (joules per mole), and i represents the species in equilibrium. The species can be ions and electrons. When two phases have common ions, the ions are the species under equilibrium; when they do not have common ions the species under equilibrium are the electrons.

For the cell given in Eq. (5) the following equilibrium relations apply

$$\mu_{e^-}^{\alpha} = \mu_{e^-}^{\beta} \tag{7}$$

$$\mu_{e^-}^{\phi} = \mu_{e^-}^{\alpha'} \tag{8}$$

Equilibrium relations between phases β and δ, and ε and ϕ are also needed. These relations can be obtained using the electrochemical reactions given in Eqs (3) and (4).

For the general electrode reaction given by [15]

$$\sum_i s_i M_i^{z_i} \longrightarrow n e^- \tag{9}$$

the equilibrium condition is given by [15]

$$\sum_i s_i \mu_i^{z_i} \longrightarrow n\mu_{e^-} \tag{10}$$

where s_i is the stoichiometric coefficient of species i in the electrode reaction, z_i is the charge number of species i, n is the number of electrodes transferred during the electrode reaction, and M_i is the compound formula or ion participating in the reaction.

Therefore, the equilibrium relation for Eq. (3) is given by

$$\mu^{\delta}_{K^+} - \mu^{\beta}_{K(s)} = \mu^{\beta}_{e^-} \tag{11}$$

and for Eq. (4) is given by

$$\mu^{\varepsilon}_{HgO(s)} + \mu^{\delta}_{H_2O} - \mu^{\phi}_{Hg(l)} - 2\mu^{\delta}_{OH^-} = 2\mu^{\phi}_{e^-} \tag{12}$$

We have specified all the equilibrium relations. The last step in the derivation is to connect the chemical equilibrium with the electrochemical potential; this can be done using the definition of electrochemical potential. The electrochemical potential of an ion was defined for first time by Guggenheim [16] as follows: the difference in the electrochemical potential of an ion between two phases is defined as the work of transferring reversibly, at constant temperature and constant volume, 1 mol from one phase to the other. Hence [15]

$$w = \mu_i^{\alpha} - \mu_i^{\beta} = z_i F(\Phi^{\alpha} - \Phi^{\beta}) \tag{13}$$

where w is the reversible work transferred (joules per mole), F is the Faraday's constant (96487 C equiv^{-1}), and Φ is the electric potential (voltage). Let us define U as the electric potential difference between phases β and α (right versus left electrode), that is

$$U = \Phi^{\beta} - \Phi^{\alpha} \tag{14}$$

where U is the open-circuit potential of the cell (voltage), and according to Eq. (13) it represents the change in the Gibbs free energy of the cell.

Applying Eqs (13) and (14) to the cell given in Eq. (5), the open-circuit potential of the cell is given by

$$FU = \mu^{\alpha}_{e^-} - \mu^{\alpha'}_{e^-} \tag{15}$$

Substituting the equilibrium relations given in Eqs (7), (8), (11), and (12) into Eq. (15) provides the equation for the open-circuit potential of the cell:

$$FU = \mu^{\delta}_{KOH} + \frac{1}{2}\mu^{\phi}_{Hg(s)} - \mu^{\beta}_{K(s)} - \frac{1}{2}\mu^{\varepsilon}_{HgO(s)} - \frac{1}{2}\mu^{\delta}_{H_2O} \tag{16}$$

where $\mu^{\delta}_{KOH} = \mu^{\delta}_{K^+} + \mu^{\delta}_{OH^-}$. In the derivation of Eq. (16), it has been assumed that the concentration in the cell is uniform. The effect of nonuniform electrolyte concentration in the cell can be included in the open-circuit potential using transport equations as shown by Newman [15].

The chemical potential is related to the concentration and activity coefficients of the species through [15]

$$\mu_i = RT \ln(m_i \gamma_i \lambda_i^{\theta}) \tag{17}$$

where m_i is the molality of species i (moles per kilogram), γ_i is the molal activity coefficient of species i, R is the universal gas constant (joules per mole kelvin), T is the cell temperature (kelvin), and λ_i^{θ} is the proportionality constant expression at a secondary reference state (kilograms per mole), which is independent of composition and electrical state, but dependent on solvent, temperature, and pressure. The use of Eq. (17) requires the definition of a reference state. The most common reference states used are

- pure solid at the cell temperature, $\gamma_i = 1$;
- pure solvent at the cell temperature, $\gamma_i = 1$.

Equation (17) can also be used for gas phase by using fugacity coefficients (ϕ_i) instead of activity coefficients (γ_i), and by using partial pressure of species P_i (atmosphere) instead of molality; then

$$\mu_i = RT \ln(P_i \phi_i \lambda_{i,g}^{\theta}) \tag{18}$$

where $\lambda_{i,g}^{\theta}$ is the proportionality constant for the gas phase (1/ atm). In this case the most convenient reference state is $\phi_i \to 1$ as $P_i \to 0$(ideal gas).

Substituting Eq. (17) into Eq. (16) and using pure solid reference state and pure solvent reference state, gives the equation for the concentration dependence of the open-circuit potential of the cell given in Eq. (5):

$$FU = FU^{\theta} + RT \ln(m_{\mathrm{KOH}} \gamma_{\mathrm{KOH}}) \tag{19}$$

where U^{θ} represents the standard open-circuit potential of the cell (voltage) and is given by

$$FU^{\theta} = \frac{1}{2}\mu_{\mathrm{Hg}}^{0} + RT \ln(\lambda_{\mathrm{KOH}}^{\theta}) - \mu_{\mathrm{K}}^{0} - \frac{1}{2}\mu_{\mathrm{HgO}}^{0} - \frac{1}{2}\mu_{\mathrm{H_2O}}^{0} \tag{20}$$

The superscript 0 indicates pure phases at the same temperature and pressure. As shown in Eq. (20), the standard potential is independent of composition but it depends on the system, solvent, temperature, and pressure. U^{θ} represents the maximum potential of the cell given at the beginning of this section (3.013 V).

We have shown how thermodynamics can be used to predict the open-circuit potential of the cell and to provide the expression for the standard open-circuit potential. Irreversible effects have not been included for cleanness; however, they can be significant in some cases. More information about irreversible effects in the cell potential can be found in the literature [15]. The final question that we need to address is, how to calculate the maximum potential of the cell using thermodynamics? We have defined that the standard open-circuit potential gives the maximum potential of the cell. This value can be calculated using tabulated thermodynamics data.

According to the definition of electrochemical potential given in Eq. (13), it does not make sense to talk about absolute potential values because only differences in potential can be measured. Values of potentials are reported and tabulated with respect to a reference electrode. The potential of the reference electrode, by definition, is zero (there is no potential difference between two electrodes of the same type). The primary reference electrode by convention is the standard hydrogen electrode (SHE) Pt/H_2 [14]

$$2H^+ + 2e^- \rightleftharpoons H_2(g) \tag{21}$$

The choice of the reference electrode depends on the system. The SHE is the common choice for aqueous electrolytes. For nonaqueous electrolytes, the choice of the reference electrode is more complex and particular for each system; for instance, for lithium-ion batteries the reference electrode is the Li/Li^+ electrode in the specific electrolyte used in the cell.

Half-cell standard potentials at 25 °C with respect to the SHE are tabulated [17–19]. The procedure explained above to determine the open-circuit potential of a cell can be used to calculate the half-cell open-circuit potential with respect to a SHE. Once again, the first step is to represent the schematic of the cells.

For the potassium electrode the half-cell schematic is given by

$$\left|\begin{matrix}\alpha\\ \mathrm{Pt/H_2(g)}\end{matrix}\right|\left|\begin{matrix}\delta\\ \mathrm{H^+/H_2O}\end{matrix}\right|\left|\begin{matrix}\beta\\ \mathrm{K_{(s)}}\end{matrix}\right|\left|\begin{matrix}\alpha'\\ \mathrm{Pt}\end{matrix}\right| \quad (22)$$

and for the mercuric oxide electrode the half-cell schematic is given by

$$\left|\begin{matrix}\alpha\\ \mathrm{Pt/H_2(g)}\end{matrix}\right|\left|\begin{matrix}\delta\\ \mathrm{H^+/H_2O}\end{matrix}\right|\left|\begin{matrix}\beta\\ \mathrm{HgO(s)}\end{matrix}\right|\left|\begin{matrix}\varepsilon\\ \mathrm{Hg(l)}\end{matrix}\right|\left|\begin{matrix}\alpha\prime\\ \mathrm{Pt}\end{matrix}\right| \quad (23)$$

Following the procedure described previously (using ideal gas as secondary reference state for the SHE) it is easy to show that for the anode half-cell reaction (Eq. 3), the standard cell potential with respect to SHE is given by

$$FU^\theta = -\mu^0_{\mathrm{K}} + \frac{1}{2}\mu^*_{\mathrm{H_2}} + RT\ln\left(\frac{\lambda^\theta_{\mathrm{K^+}}}{\lambda^\theta_{\mathrm{H^+}}}\right) \quad (24)$$

where the superscript * indicates ideal gas secondary reference state. The tabulated value for this standard cell potential is $U^\theta = 2.925$ V [14]. For the cathode half-cell reaction (Eq. 4) the standard cell potential with respect to SHE is given by

$$FU^\theta = \frac{1}{2}(\mu^*_{\mathrm{H_2}} + \mu^0_{\mathrm{HgO(s)}} + \mu^0_{\mathrm{H_2O}} - \mu^0_{\mathrm{Hg(l)}}) - RT\ln(\lambda^\theta_{\mathrm{H^+}}\lambda^\theta_{\mathrm{OH^-}}) \quad (25)$$

The tabulated value for this standard cell potential is $U^\theta = -0.098$ V [14]. Finally, the standard open-circuit potential of the cell is given by subtracting Eq. (25) from Eq. (24). This gives the same expression as Eq. (20). The value of the standard cell potential of the cell is $U^\theta = 3.023$ V.

Summarizing, the steps to calculate the open-circuit potential of cells using thermodynamics are:

- Draw schematic representation of the cell distinguishing the different phases.
- Write equilibrium relations for each of the phases (use Eqs (6) and (10)).
- Use the potential definition given by Eq. (13).
- Substitute chemical potential equilibrium relations into the potential definition.
- Express the chemical potential as a function of concentration and activity coefficients for liquid and solid phase (Eq. 17) or as a function of partial pressure and fugacity coefficients for the gas phase (Eq. 18). Define appropriate secondary reference states.
- Collect concentration dependent and independent terms.

Equations for the open-circuit potential of cells can be simplified further by defining all the activity coefficients and fugacity coefficients equal to 1. This simplification is known as *Nernst equation* for the open-circuit potential. The Nernst equation for the open-circuit potential of the cell shown in Eq. (5) is given by simplifying Eq. (19):

$$FU = FU^\theta + RT\ln(m_{\mathrm{KOH}}) \quad (26)$$

where the activity coefficient of potassium hydroxide has been approximated to 1 ($\gamma_{\mathrm{KOH}} = 1$). Eq. (26) not only neglects the effects of junction potentials (assumes constant concentration in the cell) but also neglects the interactions in the electrolyte. Cells with open-circuit potentials, which follow Nernst equation are known as *Nernstian cells*, and the ones that do not follow Nernst equation are known as *non-Nernstian cells*. Some of the reasons for a non-Nernstian cell could be that the junction potential (irreversible potential) is significant or that the interactions of the ions are strong. For instance, the open-circuit potential of the electrode LiC_6 (used

in lithium-ion batteries) with respect to a Li reference electrode does not follow Nernst equation due to the strong Li–Li interactions in the solid phase [20]. More details about the pitfalls of Nernst equation are given by Newman [15].

7.5.2 Electrode Kinetics

An important factor to analyze the performance of a battery is to characterize the electrode kinetics reaction rate, because the chemical energy is transformed into electrical energy through the electrode kinetics. The rates of the electrode reactions depend on the nature of the electrode surface, the composition of the electrolyte solution adjacent to the electrode (outside the double layer), and the electrode potential [15]. Before we analyze the typical expressions for electrode reaction rates, we need to discuss some important concepts such as double layer, and surface overpotential.

Close to the electrode surface there is a distribution of charges that is known as *double layer*. Let us suppose that we have the electrode reaction given by Eq. (3), and the electrode is immersed in a KOH/H_2O electrolyte solution. A distribution of the charges as the one shown in Fig. 8 will take place close to the electrode. The K^+ ions will tend to be close to the electrode surface while the OH^- ions will be close to the water interface due to intermolecular forces (OH^- ions will be attracted to the hydrogen atoms in the water molecules). The excess of K^+ ions close to the electrode formed the diffuse part of the double layer. The thickness of the double layer is of the order of nanometers. More details about double layer are given by others [14, 15].

The electrode potential is characterized by the surface overpotential (η_s), which is defined as the potential of the working electrode with respect to a reference electrode of the same kind placed in the solution adjacent to the electrode just outside the double layer [15]. The surface overpotential represents the deviation from the equilibrium condition of an electrode reaction given by Eq. (10); [15] therefore

$$-F\eta_s = \mu_{e^-} - \sum_i \frac{s_i}{n}\mu_i \qquad (27)$$

Expressions used for the reaction rates of the electrodes relate the current density with the surface overpotential and the concentration outside the double layer. One of the simplest relations used to express electrode reaction rates is the

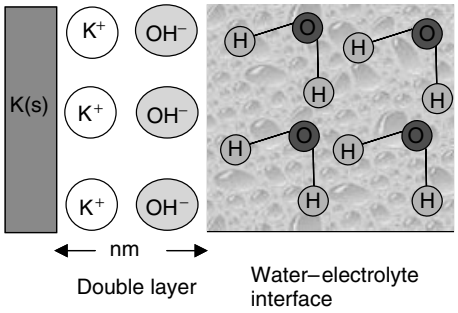


Fig. 8 Schematic representation of the double layer for the electrode reaction given in Eq. (3). The thickness of the double layer is of the order of nanometers.

Butler–Volmer equation: [15]

$$I = i_o \left[e^{\left(\frac{\alpha_a F}{RT}\eta_s\right)} - e^{\left(-\frac{\alpha_c F}{RT}\eta_s\right)} \right] \tag{28}$$

where i_o is the exchange current density (amperes per square meter) which depends on the composition, temperature, and nature of electrode surface; α_a and α_c are the anodic and cathodic transfer coefficients, respectively; and I is the current density (amperes per square meter). The transfer coefficients are the measure of the symmetry of the energy barrier of the transition state [14].

If the surface overpotential is very small ($\eta_s \to 0$), Eq. (28) can be approximated by

$$I = i_o \frac{(\alpha_a + \alpha_c)}{RT} F \eta_s \tag{29}$$

Equation (29) is known as the *linear approximation of kinetics* and is only valid for very small surface overpotentials.

If $\alpha_a F \eta_s \gg RT$ the anodic reaction is more significant than the cathodic reaction; therefore

$$I = i_o e^{\left(\frac{\alpha_a F}{RT}\eta_s\right)} \tag{30}$$

If $-\alpha_c F \eta_s \gg RT$ the cathodic reaction is more significant than the anodic reaction; therefore

$$I = -i_o e^{\left(-\frac{\alpha_c F}{RT}\eta_s\right)} \tag{31}$$

Equations (30) and (31) are known as *Tafle equations* and they are approximations of the Butler–Volmer equation, which is valid for irreversible reactions. Equation (30) is valid for irreversible anodic reactions, while Eq. (31) is valid for irreversible cathodic reactions.

The double layer can affect the distribution of the current density. When the electrode potential changes the current that flows will be distributed, part of it will be used to charge the double-layer capacity and another part will go to the faradic reaction (see Eq. 9). For a given current, the double layer will decrease the potential of the cell acting as an internal resistance (a fraction of the current will go to the double layer).

The electrode kinetics is measured experimentally. Some common measurement techniques are impedance spectroscopy, potential sweep rates, and intermittent methods. Details about these techniques are given elsewhere [14].

7.6 Transport Mechanisms in Batteries

The transport laws used in electrolytic solutions can be applied to describe the transport mechanisms in a battery. A detailed description of transport mechanisms in electrochemical systems is available elsewhere [14, 15]. Some general comments will be given in this section.

A material balance in a unit cell is given by

$$\frac{\partial C_i}{\partial t} = -\nabla N_i + \Re_i \tag{32}$$

where C_i is the concentration of species i (moles per cubic meter), t is the time (seconds), N_i is the molar flux of species i (moles per square meter second), and $\Re_i$ is the homogenous reaction rate of species i (moles per cubic meter second).

The molar flux of species i can be calculated using Nernst-Plank equation: [15]

$$N_i = -z_i u_i F C_i \nabla \Phi - D_i \nabla C_i + C_i v \tag{33}$$

where v is the velocity of the bulk (meters per second), D_i is the diffusion coefficient of species i (square meter per second), u_i is the mobility of species i (square meter mole per joule second), and $\nabla\Phi$ is the electric field potential gradient. The first term of Eq. (33) accounts for migration, the second term accounts for diffusion, and the last term accounts for convection. Migration accounts for the movements of charged species under the influence of an electric field, diffusion accounts for the movement of the species due to a chemical potential gradient (i.e. concentration gradient), and convection accounts for the movement due to stirring or hydrodynamic transport.

In a battery, and therefore in a unit cell, if the velocity of the electrolyte is very slow (no-stirring), then the convection term is usually neglected in Eq. (33) resulting in the following molar flux equation:

$$N_i = -z_i u_i F C_i \nabla\Phi - D_i \nabla C_i \qquad (34)$$

The application of Eq. (34) is limited to infinitely dilute systems (very dilute electrolytes). More details are given by Newman [15]. A more accurate representation of the flux is obtained when the gradient of the chemical potential is used for the derivation of the molar flux. This approach is known as *concentrated solution theory* [15]. For a binary electrolyte, the molar flux for the cations using concentrated solution theory becomes

$$N_+ = -\mathcal{D}\left(1 - \frac{\mathrm{d}\ln C_0}{\mathrm{d}\ln C}\right)\nabla C + \frac{I t_+^0}{z_+ v_+ F} \qquad (35)$$

where C is the electrolyte concentration (moles per cubic meter), C_0 is the solvent concentration (moles per cubic meter), $\mathcal{D}$ is the diffusion coefficient of the electrolyte (square meter per second), v_+ is the number of cations in which the specie dissociates, and t_+^0 is the transfer number of cations in the electrolyte relative to the solvent velocity. Convection has been neglected from Eq. (35).

The homogeneous reaction term, $\Re_i$ in Eq. (32), can be neglected in the separator region of the cell (assuming that the separator is chemically inert). This term can also be neglected in nonporous electrodes. When dealing with porous electrodes, the term $\Re_i$ is used to include the effect of the electrochemical reactions in the porous active material. An expression for $\Re_i$ can be obtained by using porous electrode theory. More details and references about porous electrode theory are given in the literature [15].

Other equations necessary to describe the transport of species in a unit cell or battery are the charge balance equation:

$$I = F\sum_i z_i N_i \qquad (36)$$

and the electroneutrality condition:

$$\sum_i z_i C_i = 0 \qquad (37)$$

7.7 Characteristics of Batteries

The characteristics of batteries include the values of some properties of the batteries that are necessary for the selection and use of the batteries in a specific application. The characteristics of the batteries are defined by the chemistry of the battery, that is, by the materials used for the construction of the battery (electrodes and electrolyte). The most common properties or characteristics used for the selection of a battery system are energy density, capacity, operating voltage, operating temperature, service life, cycle life (for secondary

batteries), shelf life, self-discharge rate, safety and reliability, and cost. In this section, we will describe briefly each of these characteristics or selection criteria.

The energy density is a measure of the total energy available in the battery per unit mass. Typical unit for the energy density are watt hour per kilogram (Wh/kg). The energy density is a function of the capacity and the voltage of the battery; it can be calculated as the product of the capacity and the voltage.

The capacity of a battery indicates how much current a battery is able to deliver during a 1-hour period. The common unit for the capacity is ampere hour (A h). For instance, a battery that has a capacity of 1 Ah will be discharged in 1 hour if the discharge current applied to the battery is 1 A. The capacity and the energy density of a battery pack can be modified by tuning variables such as length of the electrodes, thickness of the electrodes, particle size of the active material, porosity of the electrodes (for porous electrodes), and thickness of the separator. The criterion to increase the capacity and the energy density of a particular battery size is to pack the maximum amount of active material possible in the volume of the can (always under safe conditions, see operating temperature description). An approximate value of the capacity of the cell can be obtained from the theoretical capacity. The theoretical capacity of the cell is the maximum cell capacity and depends only on the capacity of the electrode materials. In real life, the capacity of the cell is approximately 20 to 30% lower than the theoretical capacity because of the transport and kinetics limitations inside the battery. The theoretical capacity of the cell does not account for any internal resistances inside the battery. The procedure to calculate the theoretical capacity of batteries is given in Sect. 7.8.

The operating voltage indicates the range for the potential of the cell during operation (charged and discharged values). The voltage of the battery is a function of the electrodes, electrolyte concentration, and temperature. The maximum voltage available in the battery (charged state) is given by the open-circuit potential of the battery. An approximate value of the open-circuit potential of the battery is given by the theoretical voltage. The theoretical voltage of the battery depends on the standard open-circuit potential of the electrodes (which can be calculated theoretically or measured experimentally as explained in Sect. 7.5.1). In real life, the open-circuit potential of the battery is lower than the theoretical voltage due to temperature and irreversible effects. More details about the theoretical voltage are given in Sect.7.8.

The operating temperature of the battery is very important for both the electrochemical and thermal performances of the battery. The electrochemical performance of the battery is affected by the temperature, which has influence on the transport properties, and electrochemical reaction rates. Transport properties and electrochemical reaction rates are enhanced at high temperatures and therefore internal resistances are reduced. However, high temperatures can cause chemical deterioration of the battery by enhancing secondary reactions in the electrodes and electrolytes, which may end up in a permanent capacity loss. Therefore, the electrode materials, electrolytes, separator, and so forth, should be stable at the range of temperature operation. The operating temperature also affects the thermal performance and safety of the battery through the sensible heat (heat generated because of the mass of the cell), ohmic

heat, and heat transfer to the ambience. Batteries must be designed to be safe at their temperature of operation; this is the reason why it is preferred to pack less active material in a battery can than the maximum possible amount. There is a trade-off between capacity of the battery and safety.

Service life is defined as the *length of time operation*. Service life is usually measured in hours. The service life of a battery depends on the chemistry of the battery, the type of discharge, discharge rate, capacity, and temperature. Long service lives, low discharge rates, and high temperatures (assuming the battery is chemically stable at the operating temperature) are achieved by large battery capacities. The service life represents the end of use in primary batteries and the start of a recharge cycle in secondary batteries. The end of life in secondary batteries is measured by their cycle life. A cycle is complete when the battery is discharged and charged. The cycle life of secondary batteries depends on the type of discharge, discharge rates, and temperature. The battery reaches an end due to a complete loss of capacity, which is mostly caused by undesired side reactions. The side reactions are enhanced mainly by high discharge rates and high operating temperatures.

Shelf life indicates the conditions (e.g. temperature, storage time) at which the battery can be stored. The storage conditions must be maintained carefully to reduce the self-discharge rate. Self-discharge rate is a measure of the capacity loss of a battery while it is stored. The self-discharge rate is usually measured at 20 °C, and it is indicated as the percentage loss of capacity in a year of storage (percentage loss per year). Self-discharge rates are accelerated by high storage temperatures.

Safety and reliability is a very important selection criterion for battery applications, and it considers factors such as failure rates, variability, leaking and gas generation, use of hazardous or toxic components, environmental friendly, operation under severe conditions, type of effluents, and so on.

Cost is another factor for battery selection. The cost of a battery system includes the production cost, operating or life-cycle cost, and disposing cost. The production or initial cost is affected by the design and chemistry of the battery, that is, cost of materials, type of sealing, and so on. The operating or life-cycle cost accounts for the service life and cycle life of the battery. Usually, batteries that require a high initial cost have a low operating cost. Finally, the disposing cost is the cost associated with recycling and disposal of the batteries once they are no longer useful. Battery producers are compelled to participate in a recycling and disposal program to protect the environment.

7.8 Theoretical Capacity and Voltage

The capacity and voltage of the battery was discussed in Sect. 7.7. It was mentioned that the theoretical capacity and the theoretical voltage give the maximum capacity and voltage available in the battery and that they can be used to make an estimate of these values. This section presents more details about theoretical capacity and theoretical voltage, emphasizing the procedure to calculate them.

7.8.1 Theoretical Capacity

As mentioned before, the theoretical capacity of the battery is the maximum battery capacity and depends only on

the theoretical capacity of the electrode materials. It does not account for any internal resistances inside the battery. To calculate the theoretical capacity of a battery it is necessary to calculate the theoretical capacity of the electrode materials. The theoretical capacity of the electrode materials can be calculated using Faraday's law. For the general electrode reaction given in Eq. (9), the theoretical capacity is given by

$$C_e = 26.8 \frac{\text{A - h}}{\text{eq - g}} \frac{1}{M_e} \tag{38}$$

where C_e is the theoretical capacity of the electrode (ampere hour per gram), and M_e is the equivalent weight (grams per equivalent gram, molecular weight of the electrode material divided by the number of electrons transferred during the electrode reaction) of the electrode material.

Once the theoretical capacity of both electrodes is calculated, the total capacity of each electrode should be calculated by multiplying the theoretical capacity of the electrode with the total weight of active material in the electrode, as shown

$$C_T = C_e W_e \tag{39}$$

where C_T is the total capacity of the electrode (ampere hour), and W_e is the weight of the active material in the electrode (grams).

The theoretical capacity of the cell is then given by the lowest total capacity of the two electrodes. The electrode with the lowest total capacity limits the capacity of the cell. For example, let us calculate the theoretical capacity of the Mercad cell (mercuric oxide/cadmium cell). The anodic reaction is

$$Cd + 2OH^- \longrightarrow Cd(OH)_2 + 2e^- \tag{40}$$

and the cathodic reaction is given by

$$HgO + H_2O + 2e^- \longrightarrow Hg + 2OH^- \tag{41}$$

The equivalent weight for the anode and cathode is $M_{e,Cd} = 56.201$ g/eq-g and $M_{e,HgO} = 108.295$ g/eq-g, respectively. The theoretical capacity of each of the electrodes can be calculated using Eq. (38), therefore $C_{e,Cd} = 0.477$ A - h/g and $C_{e,HgO} = 0.247$ A - h/g. Cadmium is usually in excess in a Mercad battery; therefore the capacity of the cell is limited by the capacity of the mercuric oxide electrode (positive electrode). If we use 5 g of the positive electrode, the total theoretical capacity of the battery will be $C_T = 1.235$ A - h. This means that for a discharge current of 1.235 A, the battery will be theoretically discharged in 1 hour; 1.235 A is the theoretical 1C rate of the battery; where 1C is the discharge/charge rate required to theoretically discharge/charge the battery in 1 hour.

Other authors [1] define the theoretical capacity in a different way. They define it as the total quantity of electricity involved in the electrochemical reactions, and it is usually given in ampere hour per gram. According to their definition the capacity is given by

$$C_T = \frac{1}{1/C_{e,\text{negative}} + 1/C_{e,\text{positive}}} \tag{42}$$

Equation (42) expresses the total current from both the electrodes regardless of the limiting one. The user can follow any of the two definitions for the capacity of the battery (Eqs 38 or 42); however, Eq. (38) gives a closer value to the real capacity of the battery than Eq. (42). Equation (38) is commonly used when building and testing batteries.

7.8.2 Theoretical Voltage

The theoretical voltage of the battery represents the maximum potential that can be obtained from the battery. It is measured experimentally. The theoretical voltage is the same standard potential of the battery that was defined in Sect. 7.5.1. As explained before, it can also be calculated from the standard potentials of each of the electrodes with respect to a reference electrode. For the Mercad cell, the standard potential of each of the electrodes with respect to a SHE in aqueous solution at 25 °C is $U^{\theta}_{Cd} = 0.81$ V and $U^{\theta}_{HgO} = 0.10$ V, for Eqs (40) and (41), respectively [1]. Adding both potentials gives the standard open-circuit potential of the cell, $U^{\theta} = 0.91$ V.

As explained before, the open-circuit potential of the battery depends on concentration, temperature, and transport limitations. The real voltage delivered by a battery in a closed circuit is affected by ohmic limitations (ohmic potential), concentration limitations (concentration overpotential), and surface limitations (surface overpotential). The close circuit potential of the cell is given by the open-circuit potential of the cell minus the drop in potential due to ohmic potential, concentration overpotential, and surface overpotential. The ohmic potential is due to the ohmic potential drop in the solution. It is mostly affected by the applied charge/discharge current of the battery. The concentration overpotential is associated with the concentration variations in the solution near the electrodes. It is strongly affected by transport properties such as: electrolyte conductivity, transference number, and diffusion coefficients. Finally, the surface overpotential is due to the limited rates of the electrode reactions. More details about potential limitations are given by Newman [15].

7.9 Battery Technologies

There are many different battery technologies available but only some of the most widespread ones will be discussed in this section. The literature about this topic is very extensive; only some of the most important aspects have been summarized in this section. The interested reader is referred to Refs 1, 7, 10 for more details.

7.9.1 Primary Batteries

The chemistry, advantages, disadvantages, applications, and some of the major components of the most common nonrechargeable battery technologies are summarized in this section. Table 2 presents a comparison of the energy density, open-circuit voltage, nominal voltage, end voltage, self-discharge rate, electrolyte solvents, and costs for the different primary batteries. More details about the different technologies are given next.

7.9.1.1 Leclanché's Cells

Leclanché's cells have been around for more than 100 years. They are also known as *zinc/carbon cells* or *dry cells*. They use a chemically produced manganese dioxide cathode (positive electrode), a zinc (foil or sheet) anode (negative electrode), and an aqueous electrolyte. Typical electrolyte mixtures include ammonium chloride and zinc chloride dissolved in water. The electrodes are separated by a cereal paste wet with electrolyte or a starch or polymer-coated absorbent kraft paper. The reactions

Tab. 2 Characteristics and comparison of primary batteries[a]

Characteristics		*Primary batteries*						
		Leclanché	*Magnesium*	*Alkaline manganese dioxide*	*Silver oxide*	*Zn/air*	*Li/SO_2*	*Li/MnO_2*
Energy density[b] (Wh kg^{-1})	Cylindrical shape	65	100	125	–	–	260	230
	Button shape	–	–	38	120	340	–	200
Open-circuit voltage (V)		1.5–1.75	1.9–2.0	1.5–1.6	1.6	1.45	3.1	3.5
Nominal voltage (V)		1.5	1.6	1.5	1.5	1.5	3.0	3.0
End voltage (V)		0.9	1.2	0.9	1.0	0.9	2.0	2.0
Self-discharge rate[b] (% loss $year^{-1}$)		10	3	4	6	3[c]	2	1–2
Electrolyte solvent		Aqueous	Aqueous	Aqueous	Aqueous	Aqueous	Not aqueous[d]	Not aqueous[e]
Cost (1 to 5 – least to most expensive)		1	3	2	5	3	5	4

[a] Data taken from Ref. 1.
[b] At 20 °C.
[c] If sealed.
[d] Acetonitrile.
[e] Propylene carbonate or 1,2-dimethoxyethane.

in Leclanché's cells are very complex and they have not been entirely understood yet. The basic reaction process starts with the reduction of the manganese electrode and the oxidation of the zinc electrode as follows

$$2MnO_2 + 2H_2O + 2e^- \longrightarrow 2MnOOH + 2OH^- \quad (43)$$

$$Zn \longrightarrow Zn^{+2} + 2e^- \quad (44)$$

Equations (43) and (44) take place at the positive electrode (cathode) and the negative electrode (anode), respectively. The products from the basic electrochemical reactions (Eqs 43 and 44) can react with the electrolyte solution forming different intermediate compounds depending on the discharge rate. For instance, when a solution of ammonium chloride in water is used as electrolyte, the zinc cations can react with the electrolyte according to

$$Zn^{+2} + 2OH^- + 2NH_4^+ \longrightarrow Zn(NH_3)_2^{+2} + 2H_2O \quad (45)$$

More information about the cell reactions and their mechanisms can be found elsewhere [1, 7]. For cells with ammonium chloride as electrolyte, the overall reaction for light discharge rates can be summarized by

$$Zn + 2MnO_2 + 2NH_4Cl \longrightarrow 2MnOOH + Zn(NH_3)_2Cl_2 \quad (46)$$

For cells with zinc chloride as electrolyte, the overall reaction for light discharge rates can be summarized by

$$Zn + 2MnO_2 + 2H_2O + ZnCl_2 \longrightarrow 2MnOOH + 2Zn(OH)Cl \quad (47)$$

Leclanché's cells are produced in cylindrical and flat (prismatic) designs in different sizes. Their energy density at 20 °C is very low (65 Wh kg^{-1} for the cylindrical size, see Table 2) compared to the other primary battery technologies. However, they are the cheapest batteries available in the market (see Table 2). Their nominal voltage is 1.5 V, which is comparable with other battery technologies (see Table 2).

The applications of Leclanché's cells include gas lighter, toys, watches, alarms, emergency lighting, lanterns, Geiger counters, and mass spectrometers. One of the main advantages of Leclanché's cells is that they have a long tradition of reliability; other advantages include low cell cost, large selection of shapes, and diverse sizes, voltages, and capacities. Disadvantages of these cells are low energy density, low efficiency, and poor storage life. It has also been reported that Leclanché's cells do not operate well under low-temperature conditions, they leak easily under abuse conditions, and their voltage falls steadily with discharge.

7.9.1.2 Magnesium Cells

The chemistry of Leclanché's cells was modified by changing the zinc electrode for a magnesium electrode. This change brought a new type of cells known as *magnesium cells* into the battery market. Magnesium cells use a manganese dioxide cathode (like the Leclanché's cells), a magnesium anode, and an aqueous electrolyte. Typical electrolyte mixtures include solutions of magnesium perchlorate with lithium chromate or magnesium bromide dissolved in water. Kraft paper is used as separator. The cell reactions are

$$2MnO_2 + H_2O + 2e^- \longrightarrow Mn_2O_3 + 2OH^- \quad (48)$$

$$Mg + 2OH^- \longrightarrow Mg(OH)_2 + 2e^- \quad (49)$$

Equations (48) and (49) take place at the cathode and anode, respectively. Therefore, the overall cell reaction can be summarized by

$$Mg + 2MnO_2 + H_2O \longrightarrow Mn_2O_3 + Mg(OH)_2 \quad (50)$$

Magnesium cells are produced in cylindrical shape. Their energy density at 20 °C is 100 Wh kg^{-1}, which is higher than that of Leclanché's cells but lower than that for other battery technologies (see Table 2). Their nominal voltage is 1.6 V, which is comparable with other battery technologies (see Table 2).

These cells are mainly used in military applications, such as radio transceivers, and emergency or standby equipment. The main advantage of the magnesium cells is that they double the capacity of Leclanché's cell. Other advantages of magnesium cells include good capacity retention, higher cell voltage than zinc–carbon cells, and low cost. Disadvantages of these cells are voltage delay, production of hydrogen during discharge because of the corrosion of magnesium, heat generated during use, and poor low-temperature performance.

7.9.1.3 Alkaline Manganese Dioxide Batteries

Alkaline manganese dioxide batteries have been around since 1959. They are usually known as *alkaline batteries*. They are a variation of Leclanché's cells. They use an electrolyte produced manganese dioxide as cathode (to improve its purity and increase the capacity of the cell), powdered zinc as anode (to provide a larger surface area for the reaction), and a highly concentrated solution of potassium hydroxide dissolved in water as electrolyte. Alkaline manganese dioxide batteries have a better electrochemical performance than Leclanché's cells, which is basically due to the electrolyte system and the purity and characteristics of the electrode materials used in alkaline manganese dioxide batteries. Cellulose, vinyl polymers, and polyolefins are used as separators. The cell reactions are very complex and are dependent on the discharge rate; the initial process includes the reduction of the manganese oxide electrode and the oxidation of the zinc electrode as given in

$$2MnO_2 + 2H_2O + 2e^- \longrightarrow 2MnOOH + 2OH^- \quad (51)$$

$$Zn + 2OH^- \longrightarrow Zn(OH)_2 + 2e^- \quad (52)$$

Equations (51) and (52) take place at the positive and negative electrodes, respectively. The overall cell reaction can be summarized by

$$Zn + 2MnO_2 + 2H_2O \longrightarrow Zn(OH)_2 + 2MnOOH \quad (53)$$

Alkaline manganese dioxide batteries are produced in cylindrical and button shapes. Their energy density at 20 °C is 125 and 38 Wh kg^{-1} for cylindrical and button shapes, respectively (see Table 2). Their nominal voltage is 1.5 V, which is comparable with other battery technologies (see Table 2).

The primary market of the alkaline cells is for small-scale portable applications such as cameras, toys, headphones, stereos, radios, liquid crystal TVs, and emergency lighting. Advantages of alkaline manganese dioxide batteries are good high rate discharge capability, high energy

output, good shelf life, good leakage resistance, good low-temperature performance, and good shock resistance. Their disadvantages include high initial production costs and a slightly higher operating cost than carbon-zinc cells (cylindrical type), see Table 2.

7.9.1.4 Silver Oxide Cells

Silver oxide cells were developed in the 1960s. These cells use silver oxide mixed with carbon (to increase the electronic conductivity of the material) as cathode, amalgamated pellet zinc powder as anode, and a solution of potassium hydroxide or sodium hydroxide with dissolved zincates in water as electrolyte. Permion (a radiation graft of methacrylic acid onto a polyethylene membrane) is used as separator. The cell reactions are

$$Ag_2O + H_2O + 2e^- \longrightarrow 2Ag + 2OH^- \quad (54)$$

$$Zn + 2OH^- \longrightarrow Zn(OH)_2 + 2e^- \quad (55)$$

Equations (54) and (55) take place at the cathode and anode, respectively. The overall cell reaction is given by

$$Zn + Ag_2O + H_2O \longrightarrow Zn(OH)_2 + 2Ag \quad (56)$$

Silver oxide batteries are produced in button shapes. Their energy density at 20 °C is 120 Wh kg^{-1}, which is higher than that of the alkaline manganese dioxide batteries (see Table 2). Their nominal voltage is 1.5 V, which is comparable with other battery technologies (see Table 2).

The primary market of these cells is for hearing-aid devices, photographic applications, watches, and calculators. Advantages of these cells include high energy density, high rate capability, flat discharge curve, good low-temperature performance, negligible leakage, negligible salting performance, shock and vibration resistance, and good shelf life. Their main disadvantage is their high cost, which limits the number of applications. They have one of the highest costs as indicated in Table 2.

7.9.1.5 Zinc/Air Cells

Zinc/air cells were commercialized in 1932. They use oxygen from the environment to produce electrical energy. Oxygen diffuses into the cell and is used as cathode, powdered zinc is used as anode, and an aqueous solution of potassium hydroxide is used as electrolyte. Excess of water vapor can degrade the performance of the cell; to minimize this situation a hydrophobic membrane is used as separator. The cell reactions are

$$\frac{1}{2}O_2 + H_2O + 2e^- \longrightarrow 2OH^- \quad (57)$$

$$Zn \longrightarrow Zn^{+2} + 2e^- \quad (58)$$

$$Zn^{+2} + 2OH^- \longrightarrow Zn(OH)_2 \quad (59)$$

$$Zn(OH)_2 \longrightarrow ZnO + H_2O \quad (60)$$

Equation (57) takes place at the cathode, while Eqs (58–60) take place at the anode. The overall cell reaction is given by

$$Zn + \frac{1}{2}O_2 \longrightarrow ZnO \quad (61)$$

Zinc/air batteries are produced in button shape. Their energy density at 20 °C is 340 Wh kg^{-1}, which is the highest energy density of the primary battery technologies (see Table 2); however, their applications are limited because their performance is not independent on the environmental conditions (because oxygen is obtained from the environment). Their nominal voltage is 1.5 V, which is comparable with other battery technologies (see Table 2).

Their primary market is for applications such as: man-pack radios, radar systems, electronic watches, and hearing-aid devices. Their advantages include high energy density, flat discharge voltage, long shelf life, low cost, capacity independent of temperature, and load within the operating range. Another advantage is that the cathode material does not run out from the system because it is obtained from the environment. The capacity of the battery is limited by the amount of zinc available in the cell. The cells are environmental friendly. Their main disadvantage is the limited power output. In addition, the cells have a short life once activated if the humidity of the air used is higher than the equilibrium water vapor pressure of the electrolyte mixture.

7.9.1.6 **Lithium Batteries**

Primary lithium batteries were introduced into the market in 1973, although research in this area started 20 years earlier. The major concern with the use of lithium in batteries was the stabilization and passivation of the material. Several properties make lithium a very attractive metal for battery applications, such as its light weight, high electrochemical equivalence, high voltage, and good conductivity. However, lithium is a very reactive material, especially in the presence of water. Nonaqueous electrolytes are used in lithium batteries to avoid violent reactions. Even though the electrolytes are mostly dry, lithium still reacts with the electrolyte forming a passivation layer called *solid electrolyte interface* (SEI). The formation of the SEI is convenient from a safety standpoint; it makes the material stable during the operation of the battery and provides a longer shelf life to the battery. On the other hand, the SEI causes a decrease in the capacity of the cell and it significantly modifies the electrochemical behavior of lithium. Kinetics and details about the formation of the SEI are still unknown.

Lithium primary batteries can be classified into three different categories: soluble cathode cells, solid cathode cells, and solid electrolyte cells. Their classification depends on the type of cathode material and the electrolytes used. Soluble cathode cells use liquid or gaseous cathode materials that are dissolved in the electrolyte. Because of the physical state of the cathode, they have a low resistance, which gives them a high rate capability. Nevertheless, the pressure in these cells can go really high due to the presence of gases. Solid cathode cells use solid materials as cathodes. They have the advantage of not being highly pressurized; however, they do not have the high rate capability that soluble cathode cells possess. Solid cathode cells are usually designed to handle medium discharge rates. Solid electrolyte cells use solid electrolytes, which give them an extremely long storage life, and practically eliminate pressure problems in lithium batteries. However, the internal resistance of these cells is the highest and therefore they are capable of handling only low discharge rates. The most common primary lithium batteries available in the market belong to the soluble cathode cells and solid cathode cells categories. A brief description of the most popular designs in these two categories is given next.

7.9.1.6.1 **Li/SO_2 Cells** Lithium/sulfur dioxide cells (Li/SO_2) are perhaps one of the most advanced lithium battery systems. They belong to the soluble cathode cells category. Liquid SO_2 is used as cathode; a lithium foil is used as anode, and lithium bromide dissolved in acetonitrile is used as electrolyte. The active cathode material is held on an aluminum mesh with

a carbon black/Teflon mix. Carbon black is used as filler to increase the electronic conductivity of the cathode and Teflon is used as binder. The separator is made of a microporous polypropylene. The cell reactions are

$$2SO_2 + 2e^- \longrightarrow S_2O_4^{=} \quad (62)$$

$$2Li \longrightarrow 2Li^+ + 2e^- \quad (63)$$

Equations (62) and (63) take place at the cathode and anode electrodes, respectively. The overall cell reaction is given by

$$2Li + 2SO_2 \longrightarrow Li_2S_2O_4 \quad (64)$$

Li/SO_2 batteries are produced in cylindrical shapes. Their energy density at 20 °C is 260 Wh kg^{-1}, which is the highest energy density of the primary battery technologies at the cylindrical size (see Table 2). Their nominal voltage is 3.1 V, which is twice as high as the nominal voltage of other primary batteries (see Table 2). These batteries are one of the most expensive batteries in the market.

These cells are mainly used in military equipment such as radio transceivers, portable surveillance devices, and munitions. The cells possess a low internal resistance and they can handle high current or high-power requirements. Other advantages of Li/SO_2 cells include light weight, small size, and excellent performance at low temperatures. On the other hand, these batteries may produce severe voltage delay during long storage, which is a consequence of the SEI on the surface of the anode. In addition, if the cells are abused they may explode or vent a highly toxic gas.

7.9.1.6.2 **Li/MnO_2 Cells** The lithium/manganese dioxide cells are the most popular cells produced in the solid cathode category of lithium primary batteries. Specially treated manganese dioxide is used as cathode, a lithium foil is used as anode, and a lithium salt dissolved in an organic solvent is used as electrolyte. Lithium perchlorate is one of the most common salts used in the electrolyte, while propylene carbonate and 1,2-dimethoxyethane are the main solvents used in the electrolyte mixture. Polypropylene is used as separator. The cell reactions are

$$Mn^{IV}O_2 + Li^+ + e^- \longrightarrow Mn^{III}O_2(Li^+) \quad (65)$$

$$Li \longrightarrow Li^+ + e^- \quad (66)$$

Equations (65) and (66) take place at the cathode and anode electrodes, respectively. In this case, lithium ions enter into the manganese dioxide crystal lattice; this structure is represented by $Mn^{III}O_2(Li^+)$. The overall cell reaction is summarized by

$$Li + Mn^{IV}O_2 \longrightarrow Mn^{III}O_2(Li^+) \quad (67)$$

Li/MnO_2 batteries are produced in cylindrical, flat (prismatic), and button shapes. Their energy density at 20 °C is 230 and 200 Wh kg^{-1} for cylindrical and button shapes, respectively (see Table 2). Their nominal voltage is 3.5 V, which is the highest nominal voltage of the primary battery technologies (see Table 2).

Because of the solid cathode material, these cells are used at moderate and low discharge rates. Their primary market is for memory backup applications. They are also used in watches, calculators, cameras, photoflash, motor drives, and toys. Li/manganese dioxide cells have a great performance at moderate and low discharge rates. Other advantages include high cell voltage and excellent storability even at elevated temperatures. Their major

disadvantage is the moderate and low capability rate due to the internal resistance of the cells. Because of the low discharge rate limitation, these cells are not available in large sizes.

7.9.2 Secondary Batteries

The chemistry, advantages, disadvantages, applications, and some of the major components of the most common rechargeable battery technologies are summarized in this section. Table 3 presents a comparison of the energy density, cell voltage operating range, operating temperature, self-discharge rate, cycle life, electrolyte solvents, high rate performance, and costs for the different secondary batteries. More details about the different secondary battery technologies are given next.

7.9.2.1 Lead-acid Batteries

Lead-acid batteries were developed in 1859 by Planté. Scientists including Faure, Volkmar, Brush, Gladstone, Tudor, Phillipart, and Woodward among others, contributed to improve Planté's cell and to create the lead-acid battery designs that are now available in the market. Lead-acid batteries use lead dioxide as cathode material, metallic lead as negative material, and a sulfuric acid solution as electrolyte. Different materials are used as separator, for example, microporous rubber, cellulose, polyvinyl chloride, polyethylene, and glass fiber. The cell reactions are

$$PbO_2 + H_2SO_4 + 2H^+ + 2e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} PbSO_4 + 2H_2O \quad (68)$$

$$Pb + H_2SO_4 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} PbSO_4 + 2H^+ + 2e^- \quad (69)$$

Equations (68) and (69) take place at the cathode and anode electrodes, respectively. The overall cell reaction is summarized by

$$Pb + PbO_2 + 2H_2SO_4 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2PbSO_4 + 2H_2O \quad (70)$$

Lead-acid batteries are produced in prismatic sizes. There are different models of lead-acid batteries available in the market. The models can be classified according to the application on:

- SLI (starting, lighting, ignition) batteries. These models are used in automobiles (automotive design). Their function is to start an internal combustion engine. SLI batteries discharge briefly but at high current. The most important characteristics of the SLI models are given in Table 3.
- Deep cycle and traction batteries. These models are used in electric vehicles. Traction batteries are designed to power vehicles.

The market of lead-acid batteries is divided into the following categories: automotive designs (and small energy storage designs) and industrial designs. Common applications in automotive designs (and small storage designs) are golf carts, off-road vehicles, on-road vehicles, emergency lighting, alarm signs, and sealed cells for tools, instruments, and electronic devices. Common applications in industrial designs are railway signals, switchgears, and so on.

The operating voltage of lead-acid batteries is from 2.0 to 1.8 V. Their operating voltage is higher than those for other technologies but lower than the operating voltage of lithium-ion batteries (see Table 3). Lead-acid batteries have the best

Tab. 3 Characteristics and comparison of secondary batteries[a]

Characteristics	*Secondary batteries*				
	Lead acid (SLI design)	*Nickel–Cadmium (sealed design)*	*Nickel–Hydrogen*	*Nickel–Metal hydride*	*Lithium-ion ($C/LiCoO_2$ system)*
Energy density[b] (Wh kg^{-1})	35	30–35	55	50	90
Cell voltage operating range (V)	2.0–1.8	1.25–1.0	1.3–1.15	1.25–1.10	4.0–2.5
Operating temperature (°C)	−40 to 55	−40 to 50	0 to 50	−20 to 50	−20 to 55
Self-discharge rate[b] (% loss/month)	2–3[c]	15–20	60	20	5–10
Cycle life	200–700	300–700	1500–6000	300–600	500–1000
Electrolyte solvent	Aqueous	Aqueous	Aqueous	Aqueous	Not aqueous[d]
High rate performance	Good	Good	Good	Good	Moderate
Cost (1 to 4 – least to most expensive)	1	2	4	3	3

[a] Data taken from Ref. 1.
[b] At 20 °C.
[c] Maintenance free design.
[d] Organic solvents such as propylene carbonate, ethylene carbonate, ethyl methyl carbonate, and so on.

self-discharge rate (2–3 % loss month^{-1}) of the secondary battery technologies and they are the cheapest to produce. Other advantages of lead-acid batteries include their relatively low production cost, ease of manufacture, availability of the raw materials, long cycle life, and versatile uses. Disadvantages of these batteries are poor low-temperature characteristics, limited energy density, and poor charge retention. These batteries cannot be left in the discharged state for too long without being damaged. In addition, lead-acid batteries are subject to the following self-discharge reactions:

$$PbO_2 + H^+ + HSO_4^- \longrightarrow PbSO_4 + H_2O + \frac{1}{2}O_2 \quad (71)$$

$$Pb + H^+ + HSO_4^- \longrightarrow PbSO_4 + H_2 \quad (72)$$

The self-discharge reactions cause hydrogen and oxygen evolution, which may end up causing thermal runaway in the battery. The rates of the self-discharge processes (Eqs 71 and 72) depend on temperature, electrolyte volume and concentration, and impurity content (specially antimony content). More details about the self-discharge process in these batteries is given in the literature [7].

7.9.2.2 Nickel–Cadmium Batteries

Nickel–cadmium (Ni–Cd) batteries were invented in 1899 by Jungner. They use nickel hydroxide as cathode active material, cadmium and iron as anode, and an aqueous solution of potassium hydroxide as electrolyte. Materials such as polystyrene, nonwoven nylon, and cellulose are used as separators. The cell reactions are

$$2NiO(OH) + 2H_2O + 2e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2Ni(OH)_2 + 2OH^- \quad (73)$$

$$Cd + 2OH^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} Cd(OH)_2 + 2e^- \quad (74)$$

Equations (73) and (74) take place at the cathode and anode, respectively. In practice, these reactions are more complex with the formation of different NiO(OH) modifications and a series of higher nickel oxides with different degrees of hydration. The overall reaction is summarized by

$$Cd + 2NiO(OH) + 2H_2O \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} Cd(OH)_2 + 2Ni(OH)_2 \quad (75)$$

Ni–Cd batteries are produced in prismatic, cylindrical, and button designs in different sizes. There are different kinds of Ni–Cd batteries available but all of them fall into any of the two categories: vented Ni–Cd batteries and sealed Ni–Cd batteries. Sealed Ni–Cd batteries are not designed for a pressure build up. More details about sealed Ni–Cd batteries are given in Table 3.

Ni–Cd batteries are used in different standby power applications such as: emergency lighting, emergency power source, extensive railroad, emergency brakes, and air conditioning. They can be used as a power source for signal lamps, hand lamps, and searchlights. Other applications include lightweight portable applications for mobile and military equipment. They are also used to start engines such as aircraft turbine engines and diesel engines. The different types of Ni–Cd

batteries (vented or sealed) have the following common advantages: long life cycles, reliability, low maintenance, strong and sturdy, great long-term storage, good charge retention, and they can be operated at a wide range of temperatures. The vented sintered plate has 50% higher energy density than the pocket plate. The sealed type has the advantage of being closed. Disadvantages of Ni–Cd batteries include: they are more expensive than lead-acid batteries, they require a controlled charging system to prevent thermal runaway, and they go through the ''memory effect''.

The term memory effect is used to describe a reversible process that may result in a gradual reduction of both power and capacity of the Ni–Cd batteries with cycling (it may also take place in Nickel–metal hydride batteries). The effect takes place following shallow charge–discharge cycles, that is, the effect is usually noticed when the batteries are partially discharged and recharged repetitively without the benefit of a full discharge. For instance, suppose that a Ni–Cd battery is usually discharged in three hours for a particular application when submitted to complete charge/discharge cycles; if the battery is not allowed to discharge completely before charging it, the user will notice that the next time the battery is used it will discharge in a time shorter than 3 hours. The phenomenon is called *memory effect* because the battery seems to remember the lower capacity. The total capacity of the cell is not affected during this process, because the process is reversible and can be fixed by submitting the battery to a regular full charge/discharge cycle. What appears to be happening is that some morphological change occurs in the undischarged active material, which causes the reduction of the working cell voltage during subsequent operations [7]. It is not clear what this morphological change is and how it originates; however, it has been suggested that the effect is due to an increase in the resistance of the undischarged active material. Recently, it has been suggested that the cause of the memory effect of secondary batteries that possess a nickel electrode is mainly due to the formation of γ-NiOOH at the surface of the electrode [21].

The term memory effect should not be confused with other degradation phenomena that take place in Ni–Cd batteries, which are known as *voltage depression* (reversible capacity loss) and *capacity fade* (irreversible capacity loss). Voltage depression takes place when the battery is overcharged, especially at high temperatures, which causes the formation of nickel–cadmium alloy. Nickel–cadmium alloy discharges at a lower voltage than pure cadmium, causing a reversible decrease in the capacity of the battery. The effect is reversible and can be fixed by submitting the battery to a normal full discharge/charge cycle. Irreversible capacity loss (capacity fade) is caused by a reduction of the electrolyte volume due to evaporation at high temperatures, by degradation of the separator, or by internal short circuits in the battery.

7.9.2.3 Nickel–Hydrogen Batteries

Nickel–hydrogen batteries (Ni/H_2) were developed in 1970 for aerospace applications. These batteries are a combination of the Ni–Cd technology (Ni electrode) and the fuel-cell technology (H_2 electrode). They use nickel hydroxide as cathode active material, a hydrogen electrode as anode, and an aqueous solution of potassium hydroxide as electrolyte. Asbestos (fuel-cell grade asbestos paper) and Zircar (untreated knit ZYK-15 Zircar cloth) are used

as separators. The cell reactions are

$$\mathrm{NiO(OH)} + \mathrm{H_2O} + e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Ni(OH)_2} + \mathrm{OH^-} \quad (76)$$

$$\frac{1}{2}\mathrm{H_2} + \mathrm{OH^-} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{H_2O} + e^- \quad (77)$$

Equations (76) and (77) take place at the cathode and anode, respectively. The overall cell reaction is given by

$$\frac{1}{2}\mathrm{H_2} + \mathrm{NiO(OH)} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Ni(OH)_2} \quad (78)$$

Ni/H_2 batteries are built in a particular shape, especially designed for applications such as geosynchronous earth orbit (GEO) commercial communications satellites and low earth orbit (LEO) satellites. Ni/H_2 batteries have a high energy density (55 Wh kg^{-1}) as shown in Table 3, and they have the best cycle life among secondary batteries. Another advantage of these cells is that they can tolerate overcharge. Additionally, the state of charge of the battery is correlated with the pressure of H_2, which makes it easy to have an indication of the state of charge of the battery. The specific technology used in these batteries, their design, and range of operation make these batteries the most expensive to build among secondary batteries. Another disadvantage is that their self-discharge is proportional to hydrogen pressure. Some scientists [22] have concluded that the self-discharge in Ni/H_2 batteries is due to a direct reaction of hydrogen with nickel oxide, and that the reaction rate depends linearly on hydrogen pressure indicating that the reaction is first order with respect to H_2. More details are given by Visintin et al. [22].

7.9.2.4 Nickel–Metal Hydride (Ni/MH) Batteries

Nickel–metal hydride batteries (Ni/MH) were designed after Ni/H_2 batteries. The idea was to design a battery with characteristics similar to the Ni/H_2 batteries for consumer applications. In Ni/MH batteries a metal alloy is used to store hydrogen while Ni/H_2 batteries use a direct hydrogen electrode. Ni/MH batteries were commercialized in 1990 by Sanyo and Matsushita for consumer applications. Ni/MH batteries use nickel hydroxide as active material for the cathode, a metal hydride as anode, and a solution of potassium hydroxide in water as electrolyte. The metal hydride is a type of alloy (hydrogen absorption alloy) that is capable of undergoing a reversible hydrogen absorbing/desorbing process while the battery is charged and discharged. Hydrogen absorption alloys can absorb up to a thousand times their own volume of hydrogen and they are composed of two metals, one that absorbs hydrogen exothermically and the other that absorbs hydrogen through an endothermic process. The combination of the two metals allows the release of hydrogen at normal temperature and pressure levels while the battery is operating. The most common metal hydride materials are the AB_2 series (e.g. $ZrNi_2$) and the AB_5 series (e.g. $LaNi_5$). More details about hydrogen absorption alloys can be found elsewhere [1, 7]. Separators for these batteries are made of fine fibers that allow oxygen diffusion (e.g. synthetic nonwoven material). The cell reactions are

$$\mathrm{NiO(OH)} + \mathrm{H_2O} + e^- \underset{\text{Discharge}}{\overset{\text{Charge}}{\rightleftharpoons}} \mathrm{Ni(OH)_2} + \mathrm{OH^-} \quad (79)$$

$$\mathrm{MH} + \mathrm{OH}^- \underset{\text{Discharge}}{\overset{\text{Charge}}{\rightleftharpoons}} \mathrm{M} + \mathrm{H_2O} + e^- \tag{80}$$

Equations (79) and (80) take place at the cathode and anode, respectively. The overall reaction is summarized by

$$\mathrm{MH} + \mathrm{NiO(OH)} \underset{\text{Discharge}}{\overset{\text{Charge}}{\rightleftharpoons}} \mathrm{M} + \mathrm{Ni(OH)_2} \tag{81}$$

The overall reaction does not involve the electrolyte; therefore, the electrolyte volume does not change in the battery during discharge and charge cycles, as it does in Ni/Cd batteries (see Eq. 75).

Nickel–metal hydride batteries are produced in cylindrical, button, and flat designs in different sizes. Some characteristics of these batteries are summarized in Table 3. Ni/MH batteries have a much higher energy density than Ni/Cd batteries. Their energy density is about 50 Wh kg^{-1} at 20 °C.

Applications of Ni/MH batteries include computers, camcorders, cellular phones, communication equipment, variety of cordless consumer products, high rate long cycle life applications, electric vehicles (under development), and so on. Ni/MH batteries are more environmental friendly than Ni/Cd batteries, and they are easy to dispose. Disadvantages of Ni/MH batteries include lower rate capability, poorer charge retention, and less tolerance for overcharge than Ni–Cd batteries. Like Ni/Cd batteries, Ni/MH batteries are also subject to the memory effect; a description of this phenomenon can be found in Sect. 7.9.2.2.

Ni/MH batteries are subject to self-discharge and capacity fade due to the reaction of the residual hydrogen in the cell with the positive electrode, as well as the slow decomposition of both electrodes. Self-discharge and capacity fade in Ni/MH batteries are enhanced by high temperatures.

7.9.2.5 Lithium-ion Batteries (Li-ion)

Lithium-ion batteries (Li-ion) were developed after primary lithium batteries to get rechargeable high energy density batteries. The first rechargeable batteries were made using a lithium foil and they are known as *rechargeable lithium batteries* or *secondary lithium batteries*. These batteries had a high energy density but they had a tendency to pose safety problems. Problems arose when lithium batteries were recharged, which caused lithium to deposit on the metallic lithium electrode forming a porous deposit with a much larger surface area than the original metal. The increase in the surface area produced a material more reactive and more sensitive to thermal stability than the original lithium electrode (explosions of these batteries were reported, thermal runaway). Another problem with secondary lithium batteries was the constraint to obtain a good cycle life. As mentioned in Sect. 7.9.1.6, lithium is a very reactive material that gets passivity in the electrolyte by forming a SEI. Every time the battery was recharged, lithium was deposited on the lithium electrode disturbing the SEI layer, and then new lithium from the electrode was used to rebuild the SEI layer. This process decreased the capacity of the battery and reduced its cycle life. Therefore, to obtain a reasonable cycle life, an excess of lithium was used, usually a three- to fivefold excess of lithium. These problems were solved with the development of lithium-ion batteries.

Lithium-ion batteries have been commercialized recently. In 1991, Sony introduced the first lithium-ion battery into the market. Li-ion batteries use carbon materials (which can reversibly accept and donate

significant amounts of lithium, without affecting its mechanical and electrical properties) as anodes, lithium intercalation compounds as positive electrodes, and nonaqueous solutions as electrolyte mixtures. Carbon can be used as anode material because the chemical potential of lithiated carbon is almost identical to the chemical potential of lithium. Different types of carbons are used (e.g. natural graphite, petroleum coke), and the capacity of the battery will depend on the material used as the anode. Transition metal intercalation compounds used as positive or cathode electrodes include $LiCoO_2$, $LiMn_2O_4$, and $LiNiO_2$. $LiNiO_2$ has the highest capacity, while $LiMn_2O_4$ has the lowest capacity; on the other hand, $LiMn_2O_4$ decomposes at higher temperatures than $LiNiO_2$. Again, the capacity of the battery depends on the material used as positive electrode. Electrolyte mixtures include a lithiated salt (e.g. $LiPF_6$, $LiASF_6$, $LiClO_4$) dissolved into an organic solvent (e.g. ethylene carbonate, propylene carbonate, ethyl methyl carbonate, dimethyl carbonate). Polyethylene and polypropylene are used as separators. The cell reactions are

$$x\mathrm{Li}^+ + xe^- + \mathrm{Li}_{1-x}\mathrm{AB} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{LiAB} \tag{82}$$

$$\mathrm{Li}_x\mathrm{C}_6 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} x\mathrm{Li}^+ + \mathrm{C}_6 + xe^- \tag{83}$$

Equations (82) and (83) take place at the cathode and anode, respectively. The overall cell reaction is summarized by

$$\mathrm{Li}_x\mathrm{C}_6 + \mathrm{Li}_{1-x}\mathrm{AB} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{C}_6 + \mathrm{LiAB} \tag{84}$$

Lithium-ion batteries are built in cylindrical and flat designs in different sizes. Table 3 summarizes some of the characteristics of Li-ion batteries. Li-ion batteries possess the highest energy density (90 Wh kg^{-1}) and the highest voltage range (4.0–2.5 V) among secondary batteries. Other advantages include lightweight, and no memory effect. Applications of lithium-ion batteries include cell phones, laptops, electric vehicles (under development), and satellites (under development). Disadvantages of these batteries are no high rate capability (nonaqueous electrolyte increases the internal resistance of the battery) and capacity fade. One of the mechanisms associated with capacity fade is the formation of the SEI layer; more details about this are given in Sect. 7.9.1.6. Other mechanisms associated with capacity fade are electrolyte decomposition, active material dissolution, lithium deposition, electrolyte oxidation, phase change, and so on. A summary of the different mechanisms is given by Arora et al. [3].

7.10 Mathematical Modeling of Batteries

Mathematical modeling of batteries has many applications. Once a model has been validated, it can be used to (1) predict the performance of the battery, (2) design a battery for specific applications, (3) estimate important parameters and properties (i.e. solid diffusion coefficients, activity coefficients, etc.), and (4) validate a theory proposed to explain the behavior of experimental data.

The behavior of a battery is normally hard to predict due to the complex chemical and physical processes inside the battery. A very detailed model implies the solution of the material and energy balances in the battery, and therefore the simultaneous solutions of partial

differential equations (PDE). The purpose of this section is to present a general procedure for the mathematical modeling of batteries, and to provide the reader with some references in the field. The formulation of equations required for the modeling of the battery is beyond the scope of this section. The required equations and examples of their applications can be found in some of the references cited here.

Mathematical modeling of batteries can be accomplished successfully through the following steps: (1) draw schematic diagram of the battery and define the complexity of the model, (2) formulate the equations, (3) solve model equations, (4) estimate properties and parameters, and (5) validate the model. An explanation for each of these steps is given next.

7.10.1 Schematic Diagram and Complexity of the Model

The first step is to draw a schematic diagram of the battery with a clear definition of the desired predicted variables (dependent variables) and the provided variables (independent variables). As mentioned before, a battery is a combination of several electrochemical cells connected in series or in parallel depending on the desired voltage and capacity, which increases the complexity of the model. We need to decide how the battery (set of cells) will be modeled. To do this it is necessary to define the complexity of the model. The complexity of the model depends on the purpose and the application of the model. We have classified the complexity of the model as (1) empirical models and (2) first-principle models.

One widely used approach in engineering to characterize a device is to use empirical fitting based on extensive experimental data, which has the advantage of good accuracy and easy implementation; however, the obtained model equations are cell and operation specific and cannot provide any mechanism information about the cell. Compared to empirical models, first-principle models are usually more complex and very time consuming to build, but they can be used to investigate the intrinsic mechanisms of the system and to provide reasonable explanations of some abnormal behaviors of the system. They can also be easily adjusted for different types of cells. Both of these approaches are discussed below.

7.10.1.1 Empirical Models

Empirical models use analog circuit theory to predict the performance of the battery. In this case, an appropriate circuit is used as the schematic representation of the battery. Figure 9 shows an example of a

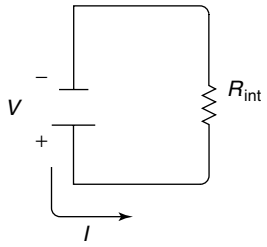


Fig. 9 Analog circuit representation of a battery.

circuit representation of a battery. This battery representation has been widely used in battery modeling. In this case, the battery performance is described by the following equation:

$$V = U + I_{app} R_{int} \tag{85}$$

where R_{int} is the internal resistance of the battery (ohms), and I_{app} is the total applied current (amperes), and V is the potential of the battery (voltage). It is assumed that the battery has a constant open-circuit potential, and the potential of a loaded battery varies linearly with the applied current. This model is actually oversimplified. It will accurately predict the potential of the battery, based on the open-circuit potential of the battery and the internal resistance. However, it cannot reflect many characteristics of a real battery, and it cannot be used for the design of batteries.

7.10.1.2 First-principle Models

First-principle modeling of batteries has become popular in recent years. It is an effective approach for battery design and optimization. Since detailed chemical and physical processes are considered, first-principle battery models are generally complicated (with a large system of differential/algebraic equations and tens or hundreds of parameters). As a result, these models usually require significant computational power and cannot be integrated easily with other simulation programs to investigate the behavior of whole power systems. In fact, most first-principle battery models are coded in FORTRAN for stand-alone simulations and are usable for only limited operating schemes.

First-principle modeling requires a detailed schematic diagram of the battery. The most common approach is to use a cell sandwich representation. An example of a cell sandwich diagram is given in Fig. 1.

7.10.2 Formulation of the Equations

Once the complexity of the model has been defined, and the appropriate schematic diagram has been drawn, a formulation of the equations is required. The discussion in this section will consider first-principle models only.

First-principle models require the formulation of the material and energy balances in the battery. The basic equations (needed to predict the electrochemical performance of the battery) are given in previous sections (see Sects 7.5 and 7.6). Equations for the energy balance are given elsewhere: Bernardi et al. [23] and Rao and Newman [24] present a description for the formulation of the energy balance in a battery. Botte et al. [25] extended Rao and Newman's energy balance description to incorporate the effect of side reactions in the thermal behavior of a cell. They also presented the details for transforming the general energy equation (Eq. 3 of Ref. 25) to the applied form of the equation (Eq. 8 of Ref. 25).

Most of the batteries use porous electrodes. Hence, the material and energy balance should consider the use of porous electrode theory. Botte et al. [26] present a review of the different approaches that have been used for porous electrode theory.

Several authors have used first-principle models to predict the performance of batteries. It is impossible to summarize all the available models within such a short account. The interested reader is referred to Refs 27–63, for examples of models for primary batteries, and to Refs 25, 26, 40,

64–90, for examples of models for secondary batteries.

7.10.3 Solution of Model Equations

Once the governing equations have been defined for the dependent variables of interest, they should be simplified and finally solved. One should check carefully for all the assumptions made in the model. Sometimes, assumptions are made implicitly; and final results do not make physical sense [26].

The governing equations can be solved using analytical or numerical techniques. Analytical techniques are the best, if available, because they show explicitly how the parameters of the system are involved. In addition, analytical techniques give much insight into a system, which is one of the primary objectives of modeling [26]. However, the equations required for battery modeling are usually very complicated and most of the time they can be solved only with numerical techniques. Botte et al. [26] present a brief review of the analytical and numerical techniques that have been used in modeling of batteries. Some of the analytical techniques are series expansion, Laplace transform, separation of variables, and perturbation methods. Some of the numerical techniques are finite difference approximation, finite element technique, control volume formulation, and method of lines. They also discuss the use of available software that help in solving the equations, programming languages, and the different trends in the solution of the governing equation that simulate the performance of the cell. The trend is to use commercially available software or many freely available robust solvers (available from NETLIB [91]) of discretized algebraic equations (DAE) to solve for the governing equations.

Recently, Wu and White [92] developed a procedure for serial simulation of cycling processes that can be used for modeling the cycling processes of rechargeable (secondary) batteries.

7.10.4 Estimation of Properties and Parameters

The parameters and properties required by the model need to be estimated or measured. The model will be as good as the parameters used. Bad input data will generate wrong results.

The measurement of transport properties, kinetics parameters, and physical properties is not an easy job. Bard and Faulkner [14] give a description of the experimental techniques that can be used to estimate electrochemical properties.

Computational chemistry and molecular simulations can also be used for the prediction of some required properties. The trend is to combine experimental techniques with computational chemistry and molecular modeling to understand and predict the physical properties of batteries [26].

7.10.5 Validation of the Model

The last step is the validation of the model with experimental data. Depending on the model, different type of data can be used for its validation, such as galvanostatic charge/discharge curves, voltammetry data, and pulse charge/discharge data.

7.11 Summary and Battery Trends

An overview about batteries was given including an understanding of how they

work, their thermodynamics, electrode kinetics, transport processes, applications, market and trends, and characteristic variables. A summary of the most common available technologies in batteries, and a description of the steps for battery modeling were also discussed. Significant changes and developments have been accomplished in the field of batteries in the last 200 years. The battery market has seen a tremendous growth in the last 10 years. It is expected that in the coming years secondary batteries will replace some of the market for primary batteries, as the production costs of secondary batteries decrease. In secondary batteries, the trend is that the production of lithium-ion batteries will take the majority of this market in the next few years due to their high energy density.

The new environmental control regulations (specially the control of gas emissions from automobiles) and the search for new alternate sources of energy make batteries an excellent choice for transportation purposes. Adapting batteries for electric vehicle applications has been very challenging for battery researchers and developers because of the need to provide electric vehicles with the same performance as conventionally powered vehicles. Nickel–metal hydride batteries and lithium-ion batteries are the most promising technologies for electric vehicle applications, because they offer technical advantages over the lead-acid batteries used in conventional vehicles and electric vehicles already in the market. However, some challenges need to be undertaken to improve these two technologies for transportation purposes; lead-acid batteries have been under study for the last 100 years, while lithium-ion batteries and nickel–metal hydride batteries have been under consideration for the last 12 years only. The challenges that need to be faced to use nickel–metal hydride batteries and lithium-ion batteries for vehicle applications include (1) cost reduction; the costs associated with the production of these batteries nowadays are very high to make them competitive with conventionally powered vehicles (part of the high production costs are due to the cost of the electrode materials, and the capital invested on research and development of these two battery technologies), (2) capacity increase, to allow the user to drive for a significant large number of miles before recharging the batteries, (3) cycling life improvement, to guarantee years of operation before replacing the batteries, and (4) improvement of high rate and thermal performance to guarantee the safety of the system, mostly in lithium-ion batteries. The key activities to accomplish these challenges are to (1) develop new electrode materials to increase the capacity of these batteries, improve their cycling rate, and reduce their production costs (both Ni/MH and Li-ion batteries), (2) to improve battery design and volume manufacturing of these batteries, so as to increase the capacity of the batteries and reduce the production costs, and (3) develop new electrolytes and additives to improve their thermal performance mostly for lithium-ion batteries. These challenges require a complete understanding of the electrode kinetics, side reactions and capacity fade mechanisms, and transport processes (including porous electrode) involved in these batteries.

Finally, the growth of battery applications in electronics, stationary sources, and transportation sources demands more research to be undertaken in the area of recycling and disposing of used batteries.

7.12 List of Symbols

AB	composite material
C	electrolyte concentration, moles per cubic meter
C_e	theoretical capacity of the electrode, amperes hour per gram
C_0	solvent concentration, moles per cubic meter
C_i	concentration of species i, moles per cubic meter
C_T	total capacity of the electrode, ampere hour
$\mathcal{D}$	diffusion coefficient of the electrolyte, square meter per second
D_i	diffusion coefficient of species i, square meter per second
F	Faraday constant, 96487 C equiv^{-1}.
i_o	exchange current density, amperes per square meter
I	current density, amperes per square meter
I_{app}	total applied current, amperes
M_e	equivalent weight, grams per equivalent gram
M_i	compound formula or ion participating in electrode reaction
m_i	molality of species i, moles per kilogram
N_i	molar flux of species i, moles per square meter second
n	number of electrodes transferred during electrode reaction
P_i	partial pressure of species i, atmosphere
R	universal gas constant, 8.314 J mol^{-1} K^{-1}
$\Re_i$	homogenous reaction rate of species i, moles per cubic meter second
R_{int}	internal resistance of the battery, ohms
s_i	stoichiometric coefficient of species i in the electrode reaction, dimensionless
T	temperature, kelvin
t	time, seconds
t_+^0	transfer number of cations in the electrolyte relative to the solvent velocity, dimensionless
U	open-circuit potential of the cell, voltage
U^θ	standard open-circuit potential of the cell, voltage
u_i	mobility of species i, square meter mole per joule second
V	potential of the battery, voltage
v	velocity of the bulk, meters per second
v_+	number of cations in which the species dissociates
W_e	weight of the active material in the electrode, grams
w	reversible work transferred in an electrochemical cell, joules per mole
x	lithium intercalation fraction, dimensionless
z_i	charge number of species i, dimensionless

Greek

Φ	electric potential, voltage
α_a	anodic transfer coefficients for Butler–Volmer equation, dimensionless
α_c	cathodic transfer coefficients for Butler–Volmer equation, dimensionless
ϕ_i	fugacity coefficient, dimensionless
γ_i	the molal activity coefficient of species i, dimensionless
η_s	surface overpotential, voltage
λ_i^θ	proportionality constant expression at secondary reference state, kilograms per mol

$\lambda_{i,g}^{\theta}$ proportionality constant for the gas phase, 1/ atm
μ chemical potential, joules per mole

Subscripts

i species in equilibrium
$+$ cation
$-$ anion

Superscripts

0 pure phases at the same temperature and pressure
* ideal gas secondary reference state

References

1. D. Linden, *Handbook of Batteries*, McGraw-Hill, New York, 1995.
2. G. G. Botte, *Mathematical Modeling and Thermal Stability of Lithium Ion Batteries*, University of South Carolina, Columbia, 2000.
3. P. Arora, R. E. White, *J. Electrochem. Soc.* **1998**, *145*, 3647–3668.
4. D. Pletcher, F. C. Walsh, *Industrial Electrochemistry*, Chapman & Hall, New York, 1990.
5. B. A. Johnson, R. E. White, *J. Power Sources* **1998**, *70*, 48–54.
6. Buchmann I. Batteries in a portable world, 2001, http://www.buchmann.ca/.
7. C. A. Vincent, B. Scrosati, *Modern Batteries*, John Wiley and Sons, New York, 1997.
8. Corrosion Doctors, 1999, http://www.corrosion-doctors.org/Batteries/History.htm.
9. Encyclopaedia Britannica Incorporation. *The New Encyclopaedia Britannica*, Encyclopaedia Britannica, Inc, Chicago, 1987.
10. R. M. Dell, *Solid State Ionics* **2000**, *134*, 139–158.
11. T. Sakai, I. Uehara, H. Ishikawa, *J. Alloys Compd.* **1999**, *293*, 762–769.
12. Japanese Market News, 2003, http://www.wtcosaka.com/market/item/batteries.html.
13. Battery Technologies Incorporation, 2006, http://www.batterytech.com.
14. A. J. Bard, L. R. Faulkner, *Electrochemical Methods Fundamentals and Applications*, John Wiley and Sons, New York, 2001.
15. J. S. Newman, *Electrochemical Systems*, Englewood Cliffs, Prentice-Hall, Inc., 1991.
16. E. A. Guggenheim, *J. Phys. Chem* **1929**, *33*, 842–849.
17. G. Milazzo, S. Caroli, *Tables of Standard Electrode Potentials*, Wiley-Interscience, New York, 1977.
18. A. J. Bard, H. Lund, *The Encyclopedia of the Electrochemistry of the Elements*, Marcel Dekker, New York, 1973.
19. A. J. Bard, J. Jordan, R. Parsons, *Standard Potentials in Aqueous Solutions*, Marcel Dekker, New York, 1985.
20. M. W. Verbrugge, B. J. Koch, *J. Electrochem. Soc.* **1996**, *143*, 600–608.
21. Y. Sato, S. Takeuchi, K. Kobayakawa, *J. Power Sources* **2001**, *93*, 20–24.
22. A. Visintin, A. Anani, S. Srinivasan et al., *J. Power Sources* **1995**, *25*, 833–840.
23. D. Bernardi, E. Pawlikowski, J. Newman, *J. Electrochem. Soc.* **1985**, *132*, 5–12.
24. L. Rao, J. Newman, *J. Electrochem. Soc.* **1997**, *144*, 2697–2704.
25. G. G. Botte, B. A. Johnson, R. E. White, *J. Electrochem. Soc.* **1999**, *146*, 914–923.
26. G. G. Botte, V. R. Subramanian, R. E. White, *Electrochim. Acta* **2000**, *45*, 2595–2609.
27. D. Bernardi, J. Newman. In *Electrochemical and Thermal Modeling of Battery, Fuel Cell and Photoenergy Conversion Systems* (Eds.: J. R. Selman, H. C. Maru), ECS, Pennington, 1986, pp. 165–192, Vol. 86–12.
28. D. Bernardi, J. Newman, *J. Electrochem. Soc.* **1987**, *134*, 1309–1318.
29. D. Bernardi, E. M. Pawlikowski, J. Newman, *J. Electrochem. Soc.* **1988**, *135*, 2922–2931.
30. H. Y. Cheh, Y. Zhang, *Tutorials in Electrochemical Engineering-Mathematical Modeling* (Eds.: R. F. Savinell, A. C. West, J. M. Fenton, J. Weidler), ECS, Pennington, 1999, pp. 141–143, Vol. 99–14.
31. J. S. Chen, H. Y. Cheh, *J. Electrochem. Soc.* **1993**, *140*, 1213–1218.
32. T. I. Evans, R. E. White. In *Electrochemical and Thermal Modeling of Battery, Fuel Cell and Photoenergy Conversion Systems* (Eds.: J. R. Selman, H. C. Maru), ECS, Pennington, 1986, pp. 22–42, Vol. 86–12.
33. T. I. Evans, R. E. White, *J. Electrochem. Soc.* **1987**, *134*, 2725–2733.
34. T. I. Evans, R. E. White, *J. Electrochem. Soc.* **1987**, *134*, 866–874.
35. T. I. Evans, R. E. White, *J. Electrochem. Soc.* **1989**, *136*, 2798–2805.

36. T. I. Evans, R. E. White, *J. Electrochem. Soc.* **1989**, *136*, 2145–2152.
37. T. I. Evans, T. V. Nguyen, R. E. White, *J. Electrochem. Soc.* **1989**, *136*, 328–339.
38. W. B. Gu, C. Y. Wang, J. W. Weidner et al., *J. Electrochem. Soc.* **2000**, *147*, 427–434.
39. M. Jain, G. Nagasubramanian, R. G. Jungst et al., *J. Electrochem. Soc.* **1999**, *146*, 4023–4030.
40. M. Jain, J. W. Weidner. in *Selected Battery Topics* (Eds.: G. Halpert, M. L. Gopikanth, K. M. Abraham, W. R. Cieslak, W. A. Adams), ECS, Pennington, 1999, pp. 202–213, Vol. 98–15.
41. M. Jain, G. Nagasubramanian, R. G. Jungst et al., in *Selected Battery Topics* (Eds.: G. Halpert, M. L. Gopikanth, K. M. Abraham, W. R. Cieslak, W. A. Adams), ECS, Pennington, 1999, pp. 162–173, Vol. 98–15.
42. J. J. Kriegsmann, H. Y. Cheh, *J. Power Sources* **1999**, *84*, 114–125.
43. J. J. Kriegsmann, H. Y. Cheh, *J. Power Sources* **1999**, *84*, 52–62.
44. J. J. Kriegsmann, H. Y. Cheh, *J. Power Sources* **2000**, *85*, 190–202.
45. M. J. Mader, R. E. White, *J. Electrochem. Soc.* **1986**, *133*, 1297–1307.
46. Z. Mao, R. E. White, *J. Electrochem. Soc.* **1992**, *139*, 1105–1114.
47. Z. Mao, R. E. White, *J. Power Sources* **1993**, *43*, 181–191.
48. J. D. Norton, C. L. Schmidt, in *Batteries for Portable Applications and Electric Vehicles* (Eds.: C. F. Holmes, A. R. Landgrebe), ECS, Pennington, 1997, pp. 389–397, Vol. 97–18.
49. E. J. Podlaha, H. Y. Cheh, *J. Electrochem. Soc.* **1994**, *141*, 28–35.
50. E. J. Podlaha, H. Y. Cheh, *J. Electrochem. Soc.* **1994**, *141*, 15–27.
51. R. Pollard, J. Newman, *J. Electrochem. Soc.* **1981**, *128*, 491–502.
52. R. Pollard, J. Newman, *J. Electrochem. Soc.* **1981**, *128*, 503–507.
53. C. L. Schmidt, P. M. Skarstad, *J. Power Sources* **1991**, *13*, 347–361.
54. C. L. Schmidt, P. M. Skarstad, *J. Power Sources* **1997**, *65*, 121–128.
55. C. L. Schmidt, A. M. Crespi, P. M. Skarstad, in *Lithium Batteries* (Eds.: R. A. Marsh, Z. Ogumi, J. Prakash, S. Surapudi), ECS, Pennington, 2000, pp. 781–789, Vol. 99–25.
56. G. D. Simpson, R. E. White, *J. Electrochem. Soc.* **1989**, *136*, 2137–2144.
57. G. D. Simpson, R. E. White, *J. Electrochem. Soc.* **1990**, *137*, 1843–1846.
58. P. M. Skarstad, C. L. Schmidt, in *Primary and Secondary Lithium Batteries* (Eds.: K. M. Abraham, M. Salomon), ECS, Pennington, 1991, pp. 85–95,, Vol. 91-3.
59. P. M. Skarstad, C. L. Schmidt, *J. Power Sources* **1993**, *43*, 111–118.
60. M. Sudoh, J. Newman, *J. Electrochem. Soc.* **1990**, *137*, 876–883.
61. B. Wu, R. Dougal, R. E. White, *J. Power Sources* **2001**, *93*, 186–200.
62. Y. Zhang, H. Y. Cheh, *J. Electrochem. Soc.* **1999**, *146*, 3566–3570.
63. Y. Zhang, H. Y. Cheh, *J. Electrochem. Soc.* **1999**, *146*, 850–856.
64. T. V. Nguyen, R. E. White, *Electrochim. Acta* **1993**, *38*, 935–945.
65. C. R. Pals, J. Newman, *J. Electrochem. Soc.* **1995**, *142*, 3282–3288.
66. C. R. Pals, J. Newman, *J. Electrochem. Soc.* **1995**, *142*, 3274–3281.
67. B. Paxton, J. Newman, *J. Electrochem. Soc.* **1997**, *144*, 3818–3831.
68. L. Song, J. W. Evans, *J. Electrochem. Soc.* **2000**, *147*, 2086–2095.
69. C. Y. Wang, W. B. Gu, B. Y. Liaw, *J. Electrochem. Soc.* **1998**, *145*, 3407–3417.
70. B. Wu, R. E. White, *J. Electrochem. Soc.* **2000**, *147*, 902–909.
71. Z. Mao, R. E. White, *J. Electrochem. Soc.* **1991**, *138*, 3354–3361.
72. J. Landfors, D. Simonsson, R. E. White, *J. Appl. Electrochem.* **1995**, *25*, 315–325.
73. J. Kim, T. V. Nguyen, R. E. White, *J. Electrochem. Soc.* **1992**, *139*, 2781–2787.
74. W. B. Gu, C. Y. Wang, *J. Electrochem. Soc.* **2000**, *147*, 2910–2922.
75. W. B. Gu, C. Y. Wang, S. M. Li et al., *Electrochim. Acta* **1999**, *44*, 4525–4541.
76. W. B. Gu, C. Y. Wang, B. Y. Liaw, *J. Electrochem. Soc.* **1998**, *145*, 3418–3427.
77. W. B. Gu, C. Y. Wang, B. Y. Liaw, *J. Electrochem. Soc.* **1997**, *144*, 2053–2061.
78. H. Gu, T. V. Nguyen, R. E. White, *J. Electrochem. Soc.* **1987**, *134*, 2953–2960.
79. T. F. Fuller, M. Doyle, J. Newman, *J. Electrochem. Soc.* **1994**, *141*, 982–990.
80. M. Doyle, J. Newman, *J. Power Sources* **1995**, *54*, 46–51.
81. M. Doyle, J. Newman, *Electrochim. Acta* **1995**, *40*, 2191–2196.
82. M. Doyle, T. F. Fuller, J. Newman, *J. Electrochem. Soc.* **1993**, *140*, 1526–1533.

83. P. De Vidts, J. Delgado, B. Wu et al., *J. Electrochem. Soc.* **1998**, *145*, 3874–3883.
84. R. Darling, J. Newman, *J. Electrochem. Soc.* **1998**, *145*, 990–998.
85. R. Darling, J. Newman, *J. Electrochem. Soc.* **1997**, *144*, 4201–4208.
86. Y. Chen, J. W. Evans, *J. Electrochem. Soc.* **1996**, *143*, 2708–2712.
87. Y. Chen, J. W. Evans, *Electrochim. Acta* **1994**, *39*, 517–526.
88. Y. Chen, J. W. Evans, *J. Electrochem. Soc.* **1994**, *141*, 2947–2955.
89. Y. Chen, J. W. Evans, *J. Electrochem. Soc.* **1993**, *140*, 1833–1838.
90. G. G. Botte, R. E. White, *J. Electrochem. Soc.* **2001**, *148*, A54–A66.
91. NETLIB, http://www.netlib.org.
92. B. Wu, R. E. White, *J. Power Sources* **2001**, *92*, 177–186.

8.1 Principles, Functions, and Classification of Fuel Cells

Marcella Cappadonia
Technische Universität München, Garching, Germany

Ulrich Stimmig
Technische Universität München, Garching, Germany

Karl V. Kordesch
Technische Universität, Graz, Graz, Austria

Julio Cesar Tambasco de Oliveira
Technische Universität, Graz, Graz, Austria

Patrik Schmuki
University of Erlangen-Nuremberg, Erlangen, Germany

8.1.1 Introduction

Energy conversion systems are the principal source of pollution on our planet. In the industrialized countries, the electric power plants which convert chemical energy into electrical energy are the main producers of CO_2 and SO_2 emissions; private cars, on the other hand, are the principal producers of CO and NO_x emissions. An alternative energy conversion technology such as the fuel cell technology offers an attractive possibility for reducing air pollution. Fuel cells convert the chemical energy of a fuel such as methane directly into electrical energy with a better yield and much lower emission levels than any other energy conversion system [1–3]. Although the principal operation of a fuel cell has been known since the last century, the necessary technology for producing fuel cell systems of high efficiency, acceptable costs, and at an industrial level has only been developed in the last few decades.

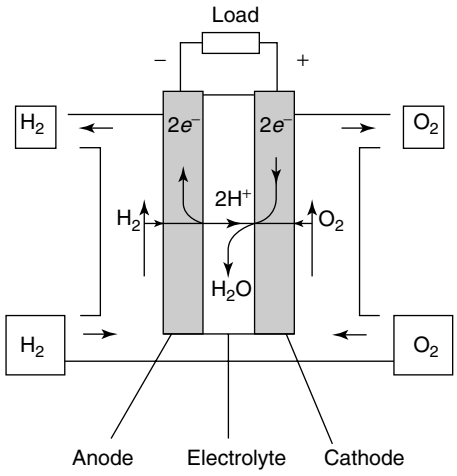
8.1.2 Principles of Fuel Cells

8.1.2.1 Functions

The principle operation of a fuel cell is comparable to that of a battery. In contrast to batteries – where the chemical energy is stored in substances inside the battery – fuel cells are just converting systems; the reagents have to be supplied continuously to the fuel cell in order to obtain electricity. Thus, fuel cells are systems which convert chemical energy directly into electricity in an invariant electrochemical setup.

Figure 1 schematically shows the principle of a fuel cell with a proton-conducting electrolyte, hydrogen as fuel and oxygen

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7



Anode:	H_2	$\rightarrow 2H^+ + 2e^-$
Cathode:	$1/2O_2 + 2H^+ + 2e^-$	$\rightarrow H_2O$
Overall reaction:	$H_2 + 1/2O_2$	$\rightarrow H_2O$

Fig. 1 Sketch of the principal operation of a fuel cell: two electrodes are separated by a gastight electrolyte. The fuel (H_2) is supplied at the anode, where it is oxidized releasing electrons. The oxidant (O_2) is supplied at the cathode, where it is reduced consuming electrons.

as oxidant. Hydrogen and oxygen are supplied to porous electrodes, the gases reach the electrode–electrolyte interface and react such that water and an electrical current are produced.

Because of the direct energy conversion, fuel cells work at a higher efficiency than energy conversion processes, which involve a conventional heating cycle. The latter are, therefore, limited by the Carnot's cycle. An obvious aspect is that fewer steps are involved in the fuel cell process than in the conventional process. In addition, the conventional process exhibits a lower conversion efficiency due to thermodynamic limitations. In a conventional process the chemical energy of fossil fuels is converted into electrical energy after several steps comprising a chemical reactor, mechanical engines, and electrical generators, with each step causing energy losses; a fuel cell can directly convert chemical energy into electrical energy, minimizing the losses.

This is illustrated in Fig. 2 which compares the theoretical efficiencies as a function of temperature of a fuel cell supplied with pure hydrogen and pure oxygen and a Carnot process as the most efficient heat cycle operating between a heat reservoir at T and one at 300 K. Only at temperatures higher than 1000 °C is

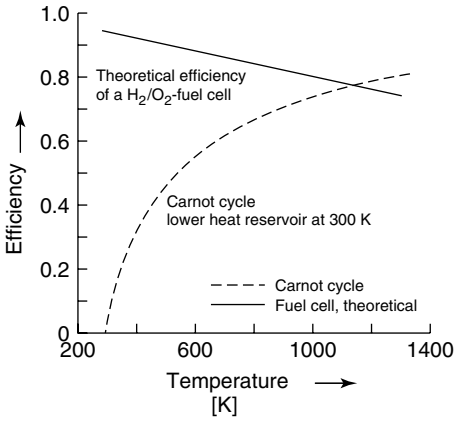


Fig. 2 Efficiency as a function of temperature of the energy conversion of a fuel cell and a conversion process limited by the Carnot's factor.

the efficiency of a Carnot process higher than that of a fuel cell. However, practical efficiencies of fuel cells remain higher even if the thermal cycles operate at much higher temperatures.

The efficiencies of the different energy conversion systems are compared in Fig. 3 as a function of the size of power plants. Figure 3 shows that the efficiencies of two types of fuel cell systems (phosphoric acid fuel cell, PAFC, see Sect. 8.1.3.1.3 and solid oxide fuel cell, SOFC, see Sect. 8.1.3.2.2) are higher than those of engines and conventional power plants of comparable size.

8.1.2.2 **Setup**

The electrodes of a fuel cell exhibit a porous structure; through the pores, fuel and oxidant reach the so-called three-phase zones, where the reagents come in contact with both the catalyst deposited on the electrodes and the electrolyte. The reactions occur in these regions. Referring to Fig. 1, the fuel (H_2) is oxidized at the anode, to form protons which move through the gas-impermeable electrolyte. At the cathode, oxygen is reduced and water is produced, while electrons flow in an external circuit. Water is the only by-product of the energy conversion in a hydrogen–oxygen fuel cell. This water is produced at the cathode in a fuel cell with proton-conducting electrolytes such as a proton exchange membrane or phosphoric acid, and at the anode in fuel cells with an alkaline electrolyte or an oxygen-ion-conducting electrolyte such as yttrium stabilized zirconia (YSZ).

Figure 1 shows only the electrochemical core of a fuel cell. However, a single electrode–electrolyte assembly is not sufficient for electricity generation, since the typical voltage of a single cell is below 1 V. In order to obtain suitable voltages, a number of cells are serially connected. Thus, in addition to the electrochemical components, a fuel cell needs interconnector plates, which connect two electrically adjacent electrochemical units and usually also

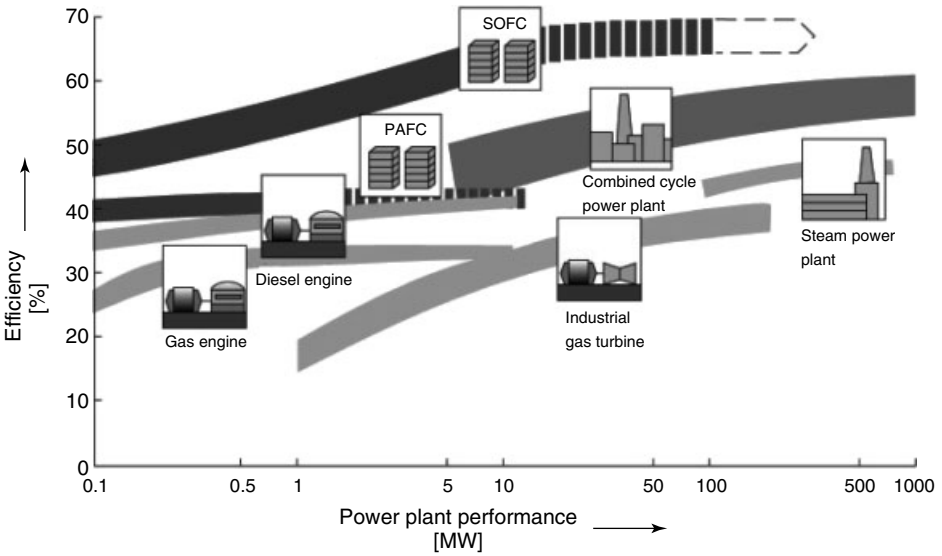


Fig. 3 Efficiency of energy conversion as a function of the size of the power plant; conventional systems are compared with projected fuel cell systems such as PAFC and SOFC.

act as the reactant distributor. Several fuel cells connected by interconnectors form a stack.

The supply of pure hydrogen and pure oxygen is not easy to realize, since no distribution system exists for these pure substances. In usual applications, a fuel cell should work with air as oxidant and with hydrogen produced in or near the system by the reforming of other fuels. Typical reactions are reforming of hydrocarbons (e.g., methane) or alcohols (e.g., methanol) (see Sect. 8.1.3.1.2). Thus, a reformer is needed in the periphery of a fuel cell system. Operation with pure chemicals involves the advantages of working at low temperature and almost problem-free operation of the catalytic materials of the electrodes, whose performance depends strongly on the presence of impurities in the reagents. In recent years many efforts have been put into the improvement of the electrocatalysts and into the gas separation process after reforming in order to use cheaper and more common fuels and air in the fuel cell. A catalytic burner for the combustion of unreacted gases and a dc/ac convertor also belong to the periphery of a fuel cell system.

The materials used in a fuel cell must exhibit properties which can be generally summarized as:

The *electrolyte* has to have a high conductivity at operating temperature, that is, it should cause a low ohmic drop ($<0.15\Omega/cm^2$). It should have negligible electronic conductivity, be gas impermeable, and chemically stable in an oxidative and reductive environment (1.2 and 0.0 V, respectively).

The *electrodes* must have a high electrocatalytic activity for fuel oxidation or oxidant reduction, a high electrical conductivity (electronic or mixed ionic and electronic), and must be chemically stable to reduction or oxidation, and gas permeable.

The *interconnector* must have a high electronic conductivity and must be stable in an oxidative environment on the cathode side and reductive environment on the anode side.

The *electrochemical system* interconnector–anode–electrolyte–cathode must exhibit high thermal stability to sufficiently withstand frequent thermal cycles, have a long operation life (5000–150 000 hours, depending on application), be light, compact, not expensive, and easy to produce. An important factor for the design is the dimension of the three-phase zones, where the electrochemical reactions take place. The three-phase zone has a thickness of just a few micrometers, therefore

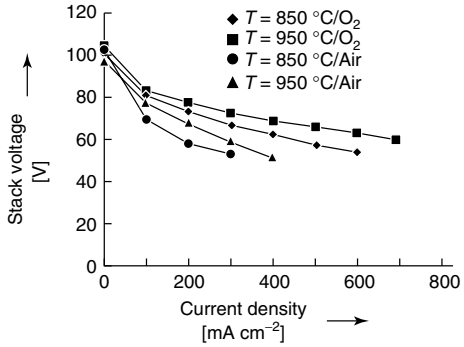


Fig. 4 Examples of current density–voltage curves for a solid oxide (see Sect. 8.1.3.2.2) fuel cell stack as a function of temperatures. These curves are measured directly. Parameter list: interconnector $CrFe_5Y$; anode Ni cermet ($Ni + ZrO_2$); cathode lanthan strontium manganite (LSM); electrolyte YSZ ($ZrO_2 - Y_2O_3$); fuel H_2 (1 bar); oxidant O_2/air (1 bar); stacks of 80 cells in serial connection.

cell designs have to consider the design of the components (electrodes, electrolytes, and interconnectors) in such a way that the specific power density is increased by minimizing the volume of the stack for the same volume of the reaction zone.

The usual way to characterize a fuel cell is to record the current density–voltage curve of single cells and stacks under different conditions. From the current density–voltage curves the current density–power curve can be obtained, which usually serves to compare different systems.

Examples of current density–voltage curves for a SOFC are given in Fig. 4.

8.1.2.3 Efficiency

Another important parameter of a fuel cell is the efficiency of the conversion of the chemicals to electricity. This *electrical efficiency* is given by the product

$$\eta_{el} = \eta_{id} \quad \eta_U \quad \eta_F \quad U_F \qquad (1)$$

where the ideal efficiency $\eta_{id} = \Delta G/\Delta H$ is the thermodynamic efficiency and corresponds to the total conversion of the free energy change, ΔG, to electrical energy in reference to the heating value of the fuel, ΔH.

The *voltage efficiency* $\eta_U = U_{cell}/U_{theoretic}$ reflects voltage losses due to polarization phenomena at the electrodes and ohmic potential drops at the interfaces in the electrolyte, and in the interconnectors (the theoretic cell voltage for water production is 1.23 V, voltages between 0.6 and 0.9 V are usually observed in real systems depending on current densities).

The *faradaic efficiency* $\eta_F = I/I_m$ is defined as the ratio between the observed current and the theoretically expected current, calculated on the basis of complete reaction. The faradaic efficiency is lower than 1 if a parallel reaction or chemical short circuit occurs in the cell.

U_F is the utilization of the fuel in the cell or stack.

The electrical efficiency of the complete fuel cell system comprises also the efficiency of the periphery and is given by

$$\eta_{el}^{system} = \eta_{el}^{stack} \quad \eta_P \qquad (2)$$

8.1.3 Classification of Fuel Cell Systems

Fuel cell systems can be divided in two groups depending on the operation temperature. Low temperature fuel cells operate at a temperature lower than 250 °C, while high-temperature fuel cells operate at a temperature above 600 °C. The fuel cell systems are usually classified and named after the electrolyte present in the cell, with one exception: the direct methanol fuel cell (DMFC) which is identified by the fuel used. Except for some alkaline fuel cells (AFCs), a common characteristic of fuel cells is the use of an immobilized electrolyte, which means that the electrolyte is solid or consists of a liquid retained on a solid matrix. Immobilized electrolytes offer many practical advantages such as fewer corrosion problems, no leakage of the liquid, easier design, planning, and construction. An overview of the material used for the components of the different fuel cell system is given in Table 1.

The other parameters such as reagents, temperature, pressure, and application of the fuel cell system depend on the electrolyte in the cell. Table 2 lists these parameters for different types of fuel cells.

Several fuels can be used; depending on the type of systems they can be directly

Tab. 1 Materials used for the components of different fuel cell systems

Fuel cell	*Electrolyte*	*Electrodes (anode/cathode)*	*Interconnector*	*Construction materials*
Alkaline fuel cell (AFC)	Concentrate KOH in porous matrix	Ni/Ag (Pt/Pt) on porous Ni	Nickel	Graphite, Inconel
Proton exchange membrane fuel cell (PEMFC)	Proton exchange membrane	Pt on graphite	Graphite sheet metal (stainless steel)	Steel
Phosphoric acid fuel cell (PAFC)	Concentrate phosphoric acid in porous matrix	Pt/Pt	Graphite	Inconel
Molten carbonate fuel cell (MCFC)	Molten Li and K carbonate in porous matrix	Ni/NiO	Inconel	Inconel, Ni, ceramic
Solid oxide fuel cell (SOFC)	Yttrium stabilized zirconia $ZrO_2(Y_2O_3)$	Ni cermet, $SrLaMnO_3$, other perovskites	Chromium based alloys or chromium based ceramics	Ceramic, Ni and Fe alloys

supplied to the anode or first undergo a reforming reaction. The thermodynamic efficiency of the combustion of different fuels are given in Fig. 5.

The temperature dependence of the theoretical efficiencies $\Delta G^\circ/\Delta H^\circ$ of various reactions are plotted on the right scale. The corresponding cell voltages calculated from the Nernst equation are shown as well (left scale).

8.1.3.1 Low Temperature Fuel Cells

8.1.3.1.1 Alkaline Fuel Cells (AFC) The first AFC was built by BACON in the mid-1950s; it was a 5 kW system with 30% aqueous KOH as electrolyte, operating at 200 °C and 5 MPa, and the reactants were pure gases. NASA used slightly modified Bacon cells for the Apollo program [4]. The technical features of the Apollo fuel cells were: 85% KOH as electrolyte, 230 °C operating temperature, 0.33 MPa pressure, pure gases as reagents, stack power 1.5 kW, and up to 11 000 hours operation at mission end. NASA improved the AFCs for the space shuttle program. The stack power was increased to 12 kW and the operating temperature lowered to 90 °C. They operated with pure gases and a specific power of 275 W kg^{-1}. Nickel plates were used as interconnectors; Pt–Pd mixtures were employed as anodic, and Au–Pt mixtures as cathodic catalysts. Lifetimes up to 15 000 hours were observed. The newest ACFs use immobilized electrolytes consisting of a porous matrix impregnated with 7 mol L^{-1} KOH; they work at temperatures <150 °C [5].

The operating principles of an AFC comprise the following reactions:

$$H_2 + 2OH^- \longrightarrow 2H_2O + 2e^- \quad \text{(oxidation of the fuel at the anode)} \qquad (3)$$

Tab. 2 Operating parameters of different types of fuel cells

Characteristic	*AFC*	*PEMFC*	*PAFC*	*MCFC*	*SOFC*
Temperature	60–90 °C	50–90 °C	160–220 °C	620–660 °C	800–1000 °C
Fuel	Pure hydrogen	Pure hydrogen, reformate[a]	Pure hydrogen, reformate	Natural gas, reformed or directly fed, biogas, coal gas	Natural gas, reformed or directly fed, biogas, coal gas
Oxidant	Pure oxygen	Pure oxygen or air	Air	Air	Air
Application	Space and military	Space, military, automotive, and stationary	Cogeneration power plant	Cogeneration or combined depending on size	Cycle power plants
System power at present		50–250 kW	11 MW	2 MW demonstration plant	100 kW demonstration plant
Electrical efficiency					
Stack (%)	69–70	50–68	55	65	60–65
System (%)	62	43–58	40	54 (cogeneration)	>50 (cogeneration)
				60–65 (combined cycle)	65–70 (combined cycle)

[a] Produced by reforming of hydrocarbons or methanol.

$$\frac{1}{2}O_2 + 2e^- + H_2O \longrightarrow 2OH^- \quad \text{(reduction of the oxidant at the cathode)} \tag{4}$$

The mobile ions are the OH^- ions of the alkaline solution, which move from the cathode to the anode. The water is produced at the anode; it can be removed together with unreacted hydrogen or with the circulating electrolyte, depending on stack design.

Reagents: The AFCs work only with pure reagents. The presence of impurities and nitrogen reduces the electrochemical efficiency of the cell. The presence of CO_2, which reacts easily with the alkaline electrolyte to form carbonates, may be detrimental for the cell, resulting in occlusion of the pores of the gas diffusion electrodes.

Electrodes: The AFCs use porous carbon electrodes. Platinum is a typical catalyst for both the anodic and cathodic reaction, but Ni and Ag are also used for the anodic and the cathodic reactions, respectively [6]. The kinetics of both reactions, but especially of oxygen reduction, is favorable in the basic environment. The charge-transfer overpotentials for the pure gases are less than 0.1 V for both electrode reactions.

Nickel is used as the material for the *interconnectors*.

The operation temperature of an AFC depends on the electrolyte concentration but it is usually below 100 °C. The possibility of working at low temperatures and the use of pure gases allows the use

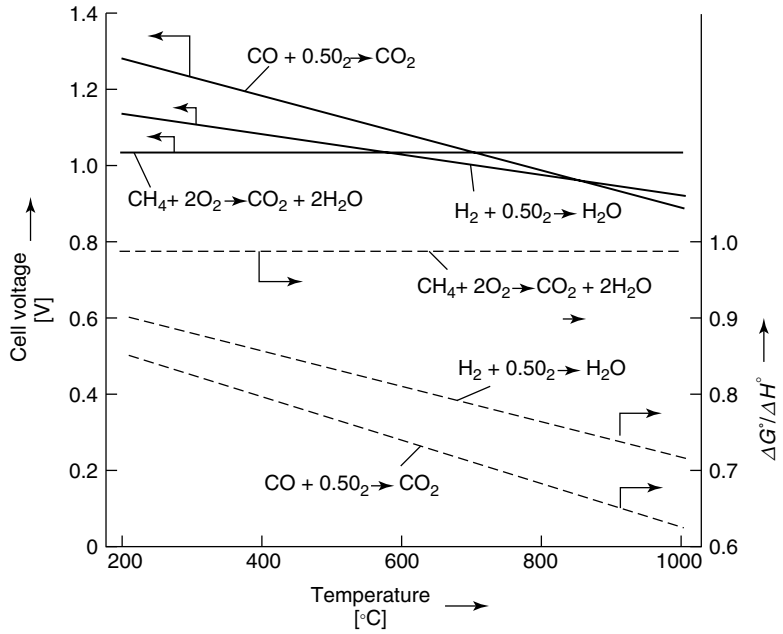


Fig. 5 Thermodynamic equilibrium voltages and theoretical efficiencies for the combustion of different fuels in a solid oxide fuel cell.

of cheaper materials for electrodes and periphery. The pure gases allow for a dead-end configuration, where no gas outlets at the electrodes are necessary.

Although AFCs exhibit the highest electrical efficiency (60%) of all fuel cell systems, their applications are limited to space and military programs, where high costs are permissible and pure hydrogen and pure oxygen can be supplied.

The small regular amounts of CO_2 contained in air are sufficient to cause problems in the operation of a AFC, thus, air must be avoided. In spite of several attempts, operation with air was not successful, since even traces of CO_2 cause carbonate formation in the pores.

8.1.3.1.2 **Proton Exchange Membrane Fuel Cells (PEMFCs)** The proton exchange membrane fuel cells (PEMFCs) are also known as *polymer electrolyte fuel cells* (PEFCs), *ion exchange membrane fuel cells* (IEMFCs), and *solid polymer electrolyte fuel cells* (SPEFCs). The first prototype of PEMFC was produced by General Electric (GE) in 1963 [7]. A 1 kW system was realized for the Gemini mission of NASA with a sulfonated polystyrene membrane as electrolyte, and the cell was operated with pure gases. In 1969, Nafion (Du Pont) was used as electrolyte in a GE fuel cell for the first time.

The operating principles of a PEMFC comprise the following reactions:

$$H_2 \longrightarrow 2H^+ + 2e^- \text{(oxidation of the fuel at the anode)} \quad (5)$$

$$\frac{1}{2}O_2 + 2e^- + 2H^+ \longrightarrow H_2O \text{ (reduction of the oxidant at the cathode)} \quad (6)$$

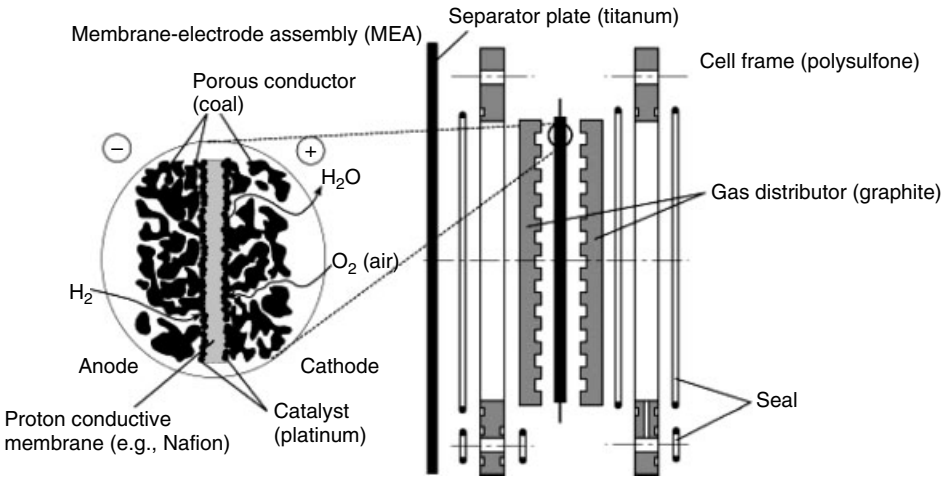


Fig. 6 Proton exchange membrane fuel cell: the membrane–electrode assembly (MEA) consists of the Nafion membrane with the electrocatalyst on the surface in contact with porous carbon electrodes. Each MEA is mounted between two gas manifolds in an electrical insulating MEA frame and it is separated by an interconnector plate from the adjacent MEAs.

The mobile ions are the H^+ which move from the anode to the cathode; the water is produced at the cathode. Water is removed together with nitrogen and unreacted oxygen, if air is used as oxidant (see Fig. 6). The PEM single cell is often called *membrane–electrode assembly* (MEA).

Reagents: Air is generally used as oxidant; pure oxygen can also be used, with the possibility of having a dead-end configuration.

Electrodes: The porous electrodes consist of carbon cloths loaded with a mixture of platinum and Nafion. In order to achieve a good contact between the electrode, the catalyst, and the electrolyte, the electrodes are pressed on the electrolyte membrane, which acts as the supporting component. Other catalysts such as binary Pt–Ru mixtures and ternary systems such as Pt–Ru–Sn have been studied for better CO tolerance [8]. The current density voltage curves of Fig. 7 show the effect of CO on the performance of a cell with platinum as catalyst. Fig. 8 shows the higher catalytic activity of Pt–Ru mixtures in the presence of CO as compared with pure Pt. The positive effect of Ru is due to its ability to provide surface oxygen at low potentials.

Graphite and stainless steel are used as materials for the *interconnectors*.

PEMFCs are characterized by the use of proton exchange membranes as electrolytes, which are good electronic insulators and very good proton conductors in the presence of water.

Nafion (DuPont) has a fluorocarbon polymer backbone with fixed sulfonic groups. Nafion membranes are produced as films with thicknesses between 0.05 and 0.18 mm and different degrees of sulfonation. They can be used as support for the porous electrodes, which can be applied by hot isostatic pressing onto the membrane [9]. A disadvantage of Nafion membranes is that their conductivity depends strongly on the water content of the membrane, which is given by the ratio of number of water molecules per available number of sulfonic acid groups (see Fig. 9) [10]. The

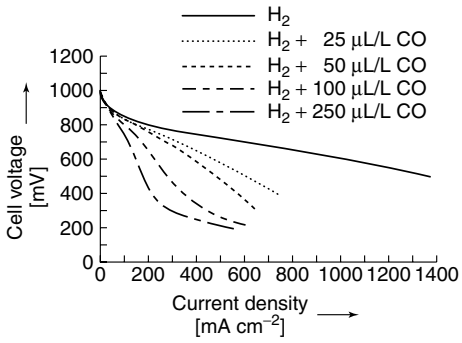


Fig. 7 Current density–voltage curves of a PEMFC with Pt as electrocatalyst as a function of CO content of the fuel; the poisoning of the electrocatalyst is evident in the reduction of the voltage of the cell at constant current density. Anode: 1 mg cm^{-2} Pt; cathode: 1 mg cm^{-2} Pt; electrolyte: Nafion 117; $p(H_2)$: 2.2 bar (abs.); $p(O_2$: 2.4 bar (abs).

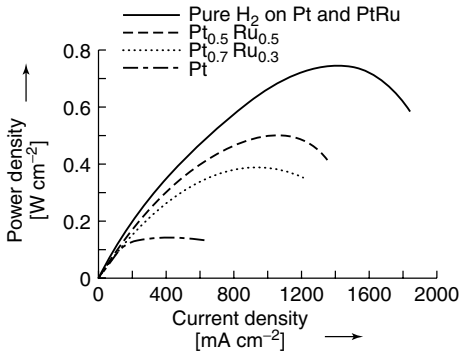


Fig. 8 Performance plots of PEM single cells with Pt, $Pt_{0.7}Ru_{0.3}$, and $Pt_{0.5}Ru_{0.5}$ anodes. The performance of different electrocatalysts in presence of CO in the fuel are shown as current density–power curves. Anode: 1 mg cm^{-2} Pt or PtRu; Cathode: 1 mg cm^{-2} Pt; Electrolyte: Nafion 117; Fuel gas: H_2 + 100 ppmv CO; Temperature: 80 °C; (pH_2): 2.2 bar (abs.); (pO_2): 2.4 bar (abs).

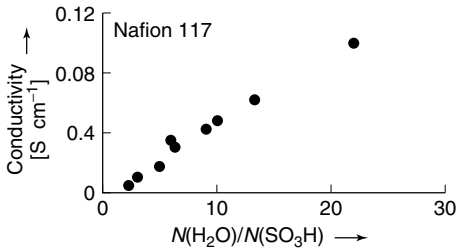


Fig. 9 Conductivity of Nafion 117 as a function of the water content of the membrane given as number of water molecules per sulfonic group. The conductivity of Nafion increases linearly with the water content.

water content of proton exchange membranes depends on the temperature, the concentration of fixed charges, and the electroosmostic drag, which causes the transport of water molecules with the proton migrating from the anode to the cathode during the operation of the fuel cell. Thus, the membrane can dry out on the anode side, causing an increase in the resistance of the electrical contact electrode–electrolyte. Waterflooding of the pores can occur at the cathode. Both processes have to be avoided by proper water management in the fuel cell, which comprises the use of humidified fuel gas.

Current densities of more than 1 A cm^{-2} at 0.5 V have been observed at PEMFC with hydrogen as fuel and air as oxidant. Even

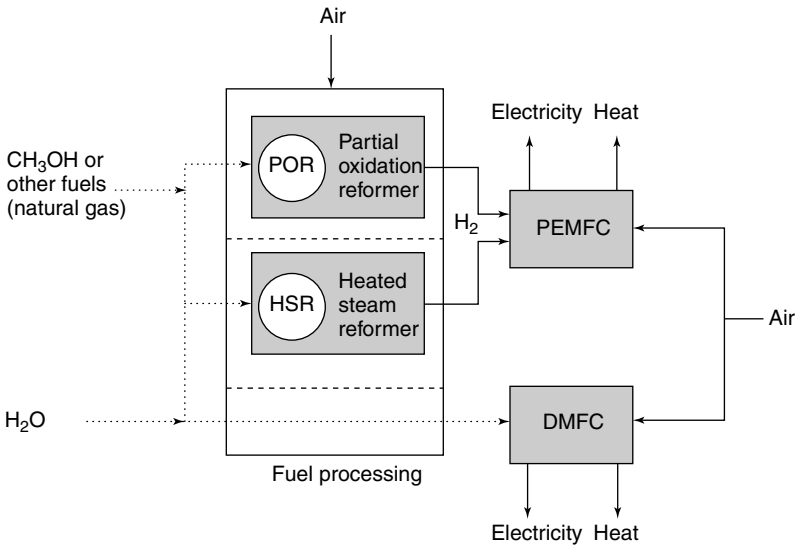


Fig. 10 Two different routes for using methanol as fuel for a fuel cell.

higher current densities can be obtained with pure oxygen as oxidant (Fig.10) [11].

The PEMFCs can also be used for *direct methanol conversion* (DMFC). In this case, the following main reactions occur in the fuel cell:

$$\text{Anode reaction } CH_3OH + H_2O \longrightarrow CO_2 + 6H^+ + 6e^- \quad (7)$$

$$\text{Cathode reaction } \frac{3}{2}O_2 + 6H^+ + 6e^- \longrightarrow 3H_2O \quad (8)$$

The oxidation occurs in several steps, with the formation of an adsorbed intermediate, such as $-CHO$ or $-CO$ on the catalyst. These adsorbates cause a deactivation of the catalyst with a consequent loss of cell efficiency. Depending on the electrolyte, a DMFC can work with a liquid methanol–water mixture at ca 80 °C or with a gaseous mixture at higher temperatures.

At the beginning of the 1980s, the interest in DMFC arose and methanol was used as fuel for high-temperature fuel cells [12]. Nowadays, PEM fuel cells with Nafion as electrolyte appear to be well suited for the direct oxidation of methanol. There are, however, some major problems in adapting a PEMFC to operate with methanol. The catalytic material of the anode has to be improved in order to avoid the loss of activity because of the formation of by-products. As in the case of a PEMFC using hydrogen from a reformer as fuel, a method to improve the anodic material is the use of Pt–Ru or Pt–Ru–Sn mixtures [8]. A particular problem of the DMFC is cross-diffusion of methanol through the electrolyte.

A possibility of avoiding the poisoning of the catalysts is to operate at a higher temperature as long as the electrolyte is stable and the cross over of methanol is tolerable. The interaction between Nafion and methanol is complex; methanol causes an increase of the ohmic drop of the

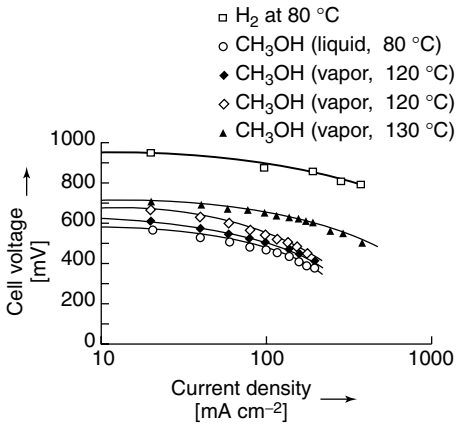


Fig. 11 Current density–voltage curves for PEMFC operating with hydrogen and methanol. The efficiency of a PEMFC operating as direct methanol fuel cell is about half of that of a PEMFC operating with pure hydrogen.

membrane and a simultaneous modification of the pore structure of Nafion. Although the PEMFC is not yet optimized for direct methanol oxidation and the high efficiency obtained in a PEM with pure hydrogen as fuel is not yet reached (see Fig. 11), current densities of 400 mA cm^{-2} at 0.5 V have been observed in a PEMFC at 130 °C for direct methanol oxidation [13].

8.1.3.1.3 Phosphoric Acid Fuel Cells (PAFC)

In contrast to alkaline electrolytes, acid electrolytes tolerate CO_2; therefore, in the 1950s, fuel cells containing acid electrolytes were proposed for the oxidation of hydrocarbons.

The operating principles of a PAFC comprise the following reactions:

$$H_2 \longrightarrow 2H^+ + 2e^- \quad \text{(oxidation of the fuel at the anode)} \quad (9)$$

$$\frac{1}{2}O_2 + 2e^- + 2H^+ \longrightarrow H_2O \quad \text{(reduction of the oxidant at the cathode, air is the most common oxidant)} \quad (10)$$

The mobile ions are the H^+ which move from the anode to the cathode through 100% phosphoric acid immobilized in a SiC–poly(tetrafluoroethylene) (PTFE) matrix. The water is produced at the cathode and is removed with nitrogen and unreacted oxygen.

Reagents: Air is used as oxidant and hydrogen, obtained by reforming of hydrocarbons such as natural gas, as fuel.

Electrodes: Graphite paper is used for the porous carbon electrodes and the catalyst is a mixture of platinum and PTFE based resin, for the use of $CO + H_2$ mixtures (reformate) Pt–Ru is used as catalyst for the anodic reaction.

Graphite is the material of the *interconnectors*.

The PAFC works at a temperature between 160 and 220 °C; the CO contamination of the catalysts is rather low at these temperatures; up to 1% CO is tolerated in the feed gas. The conversion efficiency of a complete PAFC system is ca 40%.

8.1.3.2 High-temperature Fuel Cells

8.1.3.2.1 Molten Carbonate Fuel Cells (MCFCs)

The development of molten carbonate fuel cells (MCFCs) started in 1951, when BOERS and KETLAAR built a fuel cell with a mixture of sodium carbonate and tungsten oxide as electrolyte; later they changed it to a mixture of lithium,

sodium, and potassium carbonate in a sintered MgO matrix. Such mixtures exhibit high ionic conductivities at about 700 °C.

The operating principles of an MCFC comprise the following reactions:

$$H_2 + CO_3^{2-} \longrightarrow H_2O + CO_2 + 2e^- \quad \text{(oxidation of the fuel at the anode)} \quad (11)$$

$$\frac{1}{2}O_2 + CO_2 + 2e^- \longrightarrow CO_3^{2-} \quad \text{(reduction of the oxidant at the cathode; air is the most common oxidant)} \quad (12)$$

The mobile ions are CO_3^{2-} ions which move from the cathode to the anode through the molten Li, K, and Na carbonates immobilized in a MgO matrix; the water is produced at the anode. Water is removed together with CO_2, possible traces of CO react at the anode with H_2O, forming CO_2 and H_2. Therefore, it is very convenient to use MCFC with fuels such as natural gas or other hydrocarbons.

Reagents: Hydrogen, reformed alcohols, or hydrocarbons are used as fuel and oxygen or air as oxidant.

Electrodes: The electrocatalytic material of an MCFC is nickel. The cathode becomes oxidized and lithiated during the first hours of the operation. Nickel oxide is soluble in molten carbonates; thus in the course of the operation two undesirable effects may occur: (1) metallic nickel particles are formed in the electrolyte which can lead to an electronic short circuit of the electrodes, (2) the size of the cathode diminishes. Two approaches have been proposed for solving these problems: the use of less corrosive molten carbonate mixtures and more stable cathodes containing iron and cobalt. The nickel anodes usually contain chromium, which promotes the sintering process.

Inconel is used for the *interconnectors*.

The MCFCs work at ca 650 °C.

MCFCs do not need external reforming of methane, as reforming takes place at the anode of the fuel cell. The reactions occurring at the anode are as follows (see Fig. 12):

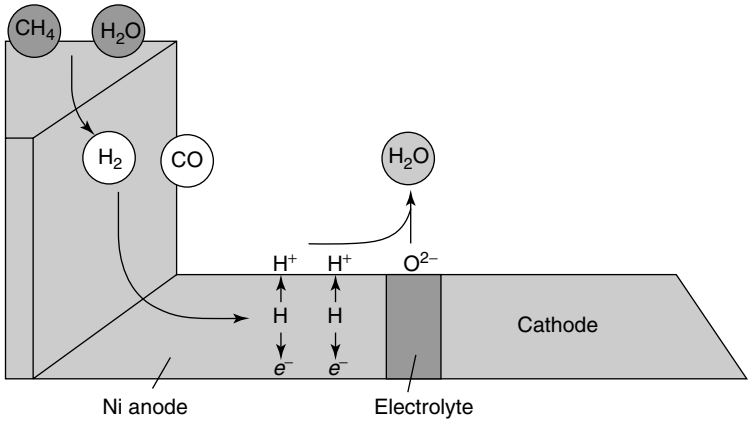


Fig. 12 Methane reforming reaction at the anode of a molten carbonate fuel cell. The reaction between fuel and water takes place in the outer part of the anode, the produced hydrogen reacts at the interface anode–electrolyte.

$$CH_4 + H_2O \longrightarrow 3H_2 + CO \quad (13)$$

$$3H_2 + 2CO_3^{2-} \longrightarrow 3H_2O + 3CO_2 + 6e^- \quad (14)$$

$$CO + CO_3^{2-} \longrightarrow 2CO_2 + 2e^- \quad \text{(minor side reaction)} \quad (15)$$

The internal reforming takes place at temperatures of about 650 °C, which is lower than the usual temperature of a reformer (about 800 °C). Advantages and problems of internal reforming of higher temperature fuel cells in general are summarized below:

Advantages: The efficiency of the power plant increases with the increasing degree of internal methane reforming from 2 at 50% internal reforming to 8 at 99% internal reforming.

Disadvantages: The H_2 producing endothermic reforming reaction is much faster than the H_2 consuming electrochemical oxidation; temperature gradients cause material problems; a sharp cooling effect occurs at the cell inlet.

8.1.3.2.2 **Solid Oxide Fuel Cells (SOFC)**

The most common solid oxide, which is used as electrolyte in SOFC, is a mixture of zirconia and yttria. YSZ contains between 8 and 10 mol% yttrium. It is an oxide ion conductor with a sufficiently high conductivity at temperatures around 1000 °C. This is an advantage for the direct oxidation of fuels in the fuel cell, but is a very challenging temperature for any material. Thermodynamics predict an efficiency of 80%. A further advantage of the high operation temperature is the possibility to efficiently use the waste heat in a combined cycle power plant.

The operating principles of an SOFC comprise the following reactions:

$$H_2 + O^{2-} \longrightarrow H_2O + 2e^- \quad \text{(oxidation of the fuel at the anode)} \quad (16)$$

$$\frac{1}{2}O_2 + 2e^- \longrightarrow O^{2-} \quad \text{(reduction of the oxidant at the cathode)} \quad (17)$$

The mobile O^{2-} ions in YSZ move from the cathode to the anode, the water is produced at the anode, and it can be removed together with unreacted hydrogen and CO_2 (in the case of hydrocarbon combustion). For internal reforming the anode reaction also provides the necessary water for steam reforming. The operation principles of a SOFC with YSZ as electrolyte is shown in Fig. 13.

Reagents: Any kind of fuels and oxidants are suitable, no reformer is needed for methane, the reagents can be directly supplied to the cell.

Electrodes: The anodes of SOFC consist of Ni cermet, a composite of metallic Ni and YSZ, Ni provides the high electrical conductivity and catalytic activity, zirconia provides the mechanical, thermal, and chemical stability. In addition, it confers to the anode the same expansion coefficient of the electrolyte and renders compatible anode and electrolyte. The electrical conductivity of such anodes is predominantly electronic. Figure 14 shows the three-phase boundary at the interface porous anode YSZ and the reactions which take place. The cathode of the SOFC consists of mixed conductive oxides with perovskite crystalline structure. Sr doped lanthanum manganite is mostly used, it is a good *p*-type conductor and can contain noble metals.

Interconnectors: A ceramic consisting of lanthanum chromite or a chromium based

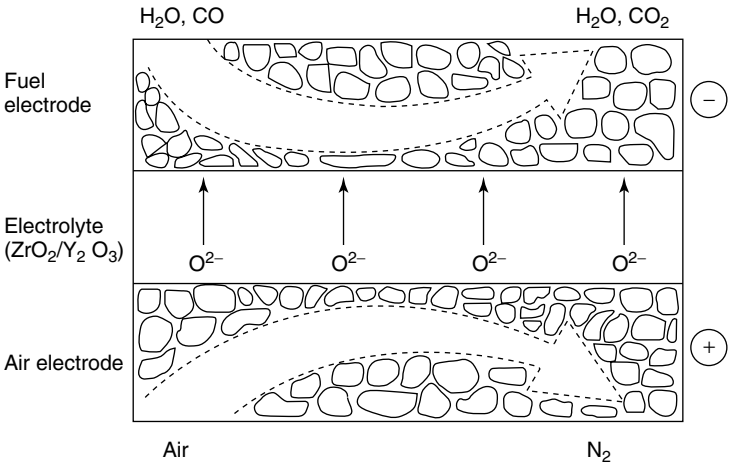
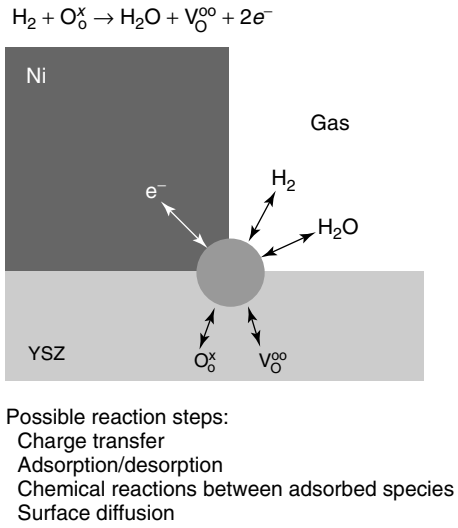


Fig. 13 Structure of the three-phase zones at anode and cathode of a solid oxide fuel cell.

Fig. 14 Anodic reactions at the three-phase boundary of a SOFC V_0^{oo} = oxygen vacancy in solid conducting oxide.



alloy are normally used; austenitic steel is also suitable at $T < 850\,^{\circ}C$.

Only a few materials can sustain the working conditions of SOFC without significant changes after long term operation. Most of the materials used in SOFC are relatively expensive. The expected high efficiency of a SOFC is quite attractive and many cell designs have been proposed. Many attempts have also been made to lower the operating temperature in order to reduce the demands on the materials.

Three designs for SOFC stacks are described below:

$$H_2 + O_0^x \longrightarrow H_2O + V_0^{oo} + 2e^-$$

Tubular Design of Westinghouse. Figure 15 shows the principles of the tubular design developed by Westinghouse. A single cell has the form of a long tube. The tube consists of a porous ceramic substrate, usually on zirconia basis. The components of the electrochemical cell are deposited on the tubular support, starting with the cathode. The electrolyte is deposited on the cathode, but the cathode is not completely covered by the electrolyte material, a strip of cathodic material is covered by the interconnector material for

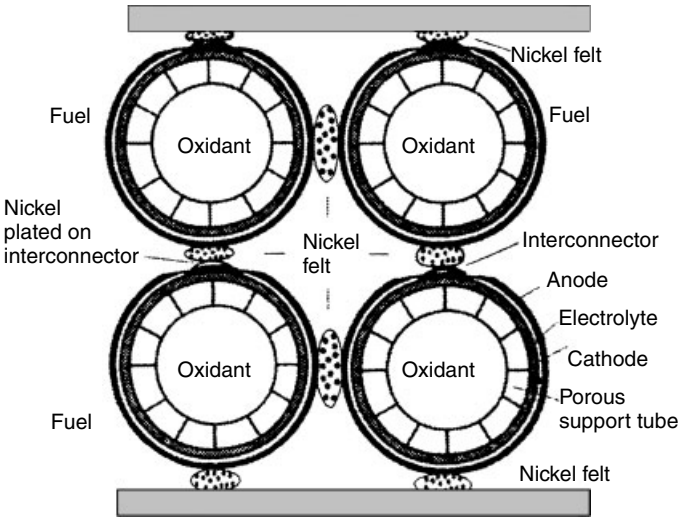


Fig. 15 Tubular SOFC stack concept of Westinghouse. Advantages: no sealing; disadvantages: ohmic loses, low power density.

the electrical connection of the cathode. The anode is deposited on the electrolyte.

Several tubes are arranged in a case separated by nickel felts, air is supplied inside the tubes, the fuel is supplied outside the tubes.

The main advantage of the tubular design is the absence of sealing in the stack. Westinghouse built a 25 kW stack and tested single cells up to 40 000 hours.

Planar Designs

Siemens Concept (Fig. 16). The planar concept of SOFC developed by Siemens is similar to the sandwich structure of the PEM single cell (see Fig. 6). A foil of YSZ with a thickness of 150–200 μm acts as the supporting component on which the electrode materials are deposited; the single cells are then stacked with interconnectors with the gas channels and sealing elements in between. The gases can be supplied in co-, cross-, or counterflow and can be let to and out of the stack in internal or external manifolding. The differences in the gas management cause different temperature patterns in the single cells and in the stack. The thickness of the electrolyte, which is defined by the necessary mechanical stability, causes ohmic losses and renders it difficult to reduce the operating temperature considerably below 900 °C. Another disadvantage of this type of planar cell is the relative small size of the single cell (commonly 5×5 cm^2, 10×10 cm^2 possible), which is mainly due to the difficulty to cast large uniform ceramic foils and the necessity to increase the thickness of the electrolyte in the case of larger cells, which results in higher ohmic losses.

IEV Concept (Fig. 16). In the planar SOFC developed by IEV-Forschungszentrum Jülich the anode is chosen as the supporting element [14]. In this cell, the anode has a thickness of approximately 1 mm, and the electrolyte and cathode are deposited in 2–10 μm and 40–50 μm thick layers, respectively,

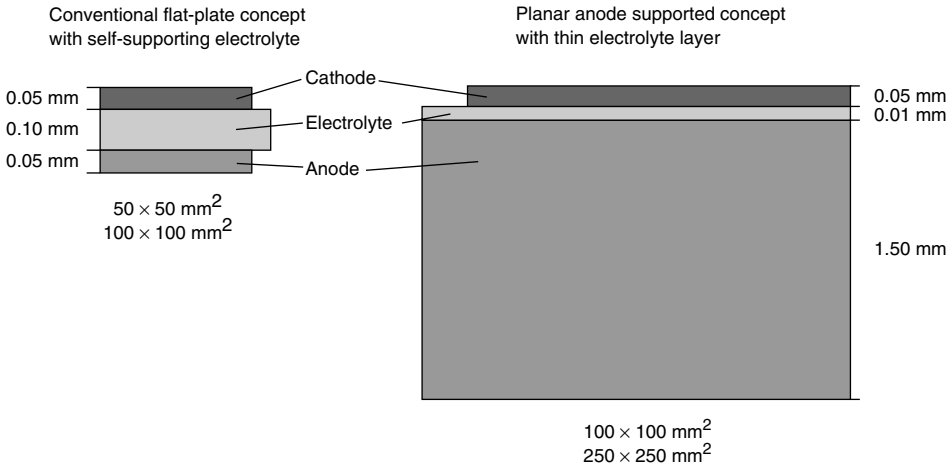


Fig. 16 Conventional flat-plate (Siemens, left) and planar (Research Center Jülich, Germany, right) SOFC stack concepts.

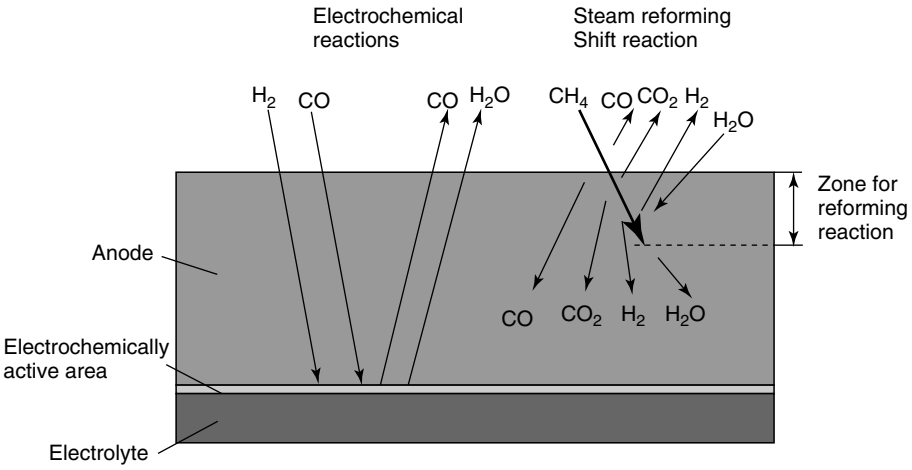


Fig. 17 Model of the multizones of a substrate anode.

on the anode. The anode exhibits a varying composition. Typical characteristics of the anode are (Fig. 17): the reforming reactions occur in the outer layer of the anode, the central part of the anode has a high electronic conductivity and high porosity, and the electrochemical reactions take place as usual at the interface with the electrolyte. This planar concept overcomes the disadvantages of the first planar concept (Siemens); it is, in fact, possible to work at lower temperature ($<700\,^{\circ}C$) and to produce large single cells ($25 \times 25\ cm^2$). In addition, the reaction zones for reforming methane and electrochemically oxidizing the resulting hydrogen are spatially separated and can be optimized separately.

References

1. M. Cappadonia, U. Stimming, K. V. Kordesch et al., *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, 2003, Vol. 15, pp. 95–117.
2. K. Kordesch, G. Simader, *Fuel Cells*, VCH, Verlagsgesellschaft, Weinheim, 1996.
3. K. Ledjeff (Ed.), *Brennstoffzellen*, C.F. Müller, Heidelberg, 1995.
4. R. B. Ferguson, Pratt & Whitney Aircraft, Design and Development of H_2/O_2-Fuel cell power plants for Apollo Command Module, NASA/Houston Final report 6/16/1962, Apollo Fuel Cell System, Proceedings of 23rd Annual Power Sources Conference, 1969, pp. 11–13.
5. M. Schautz, G. Dudley, Testing of Space fuel cells at ESTEC reprinted article from ESA Bulletin no. 60, 1992.
6. O. Führer et al., *Int. J. Hydrogen Energy* **1994**, *19*(3), 343–348.
7. R. Cohen, Gemini fuel cell system, Proceedings 20th Power Sources Conference, 24-26 May 1996, pp. 21–24.
8. A. Aramata, M. Masuda, *J. Electrochem. Soc.* **1991**, *138* 1949.
9. Brochure DuPont Nafion perfluorinated membranes, DuPont, Fayetteville, 1994.
10. T. E. Springer, T. A. Zawodzinski, S. Gottesfeld, *J. Electrochem. Soc.* **1991**, *138* 2334.
11. T. E. Springer, M. S. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1993**, *140* 3513.
12. D. J. Wheeler et al., *Proceedings Workshop on Direct Methanol – Air Fuel Cells*, The Electrochemical Society, Pennington, 1992, p. 193, Vol. 92-14.
13. M. Waidhas et al., Research and development of low temperature fuel cells at Siemens, 1994-Fuel Cell Seminar, San Diego, 1994, p. 477.
14. L. G. J. Haart, J. Divisek, E. Riensche, U. Stimming, EP-A 1 0 829 103, 1998.

8.2
Physical Modeling of Fuel Cells and their Components

Michael Eikerling
Simon Fraser University, Burnaby, BC, V5A 1S6, Canada and Institute for Fuel Cell Innovation, National Research Council Canada, Vancouver, BC, V6T 1W5, Canada

Alexei A. Kornyshev
Imperial College London, London SW7 2AZ, United Kingdom

Andrei A. Kulikovsky
Institute for Materials and Processes in Energy Systems III, Research Center Jülich, D-52425 Jülich, Germany

8.2.1
Introduction

Polymer-electrolyte fuel cells (PEFC and DMFC) possess a exceptionally diverse range of applications, since they exhibit high thermodynamic efficiency, low emission levels, relative ease of implementation into existing infrastructures and variability in system size and layout. Their key components are a proton-conducting polymer-electrolyte membrane (PEM) and two composite electrodes backed up by electronically conducting porous transport layers and flow fields, as shown schematically in Fig. 1(a).

The state-of-the-art proton-conducting polymer membranes use water networks to conduct protons. Thus, they require sufficient membrane hydration for functioning, which limits operation of the present-day cells to temperatures lower than 100 °C. A number of key points where improvement is vitally important are related to this constraint. These are: maintaining the fuel cell water balance (avoiding local dehydration of membrane or anode as well as flooding on the cathode side), improving electrode kinetics, providing large effective catalyst areas, maximizing catalyst utilization, avoiding anode catalyst poisoning due to carbon monoxide in the fuel stream and minimizing mass transport limitations in membranes, electrodes, and supporting structures.

The outstanding perspectives of fuel cells have spurred the huge economic interest in this upcoming technology, generating a highly competitive environment for industrial and academic activities. This background creates a rapidly evolving, multifaceted setting for research, dealing broadly with the engineering aspects of system optimization and materials research. The former category involves fabrication and assembly processes, operation conditions, and system or power plant layout. The latter pursues the development of new materials for charge transport, reactant and product transport, and reactivity, with structure and composition optimized to provide the best technical and economical characteristics of the cells.

Altogether, the road toward commercially available fuel cells is a subject of complex energy process engineering and material science. Smart engineering needs modeling, which parameterizes competing processes in cells and stacks in its wealth. In general, the modeling is aimed to bridge the gap between generic understanding of underlying physical and chemical processes in fuel cells, on the one hand, and performance evaluation under varying operating conditions, on the other. Modeling receives parameters of processes

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

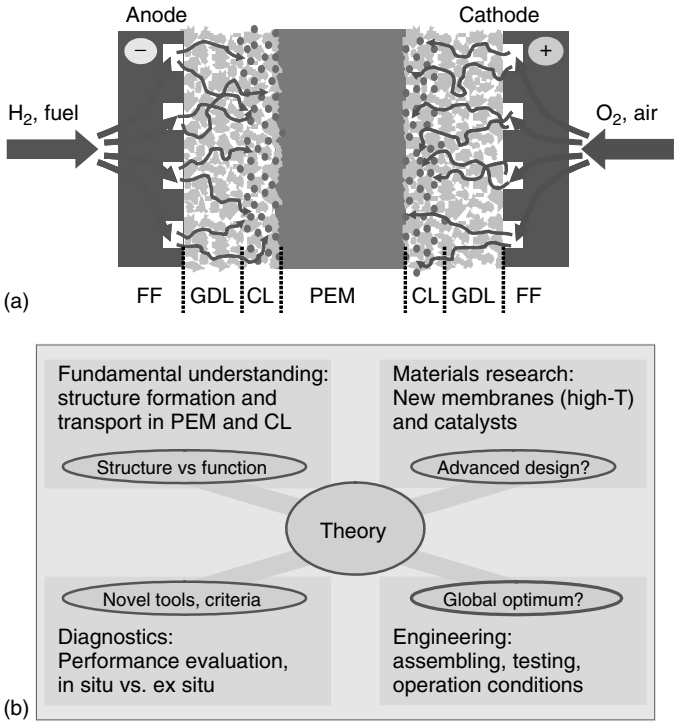


Fig. 1 (a) Principal layout of a PEM fuel cell with the main functional components, viz. proton-conducting polymer-electrolyte membrane (PEM), catalyst layers on anode (ACL) and cathode sides (CCL), gas-diffusion layers (GDL) and flow fields (FF). (b) Disciplines in fuel cell research and how they are connected by the theory.

involved in fuel cell operation from proper fundamental studies of these processes, performed ex situ or for idealized model systems. Lessons from the modeling can be passed on, thereafter, for purposes of diagnostics, materials research, and cell and stack design or they can serve as modules for larger scale system simulations.

Evidently, modeling of fuel cells is a delicate task, which comprises a multitude of functions, ranging from posing problems for research programs in fundamental science to laying out routes toward optimization of fuel cell performance. Relevant areas of fundamental research are routed in surface science, heterogeneous catalysis, electrocatalysis, nanoparticle science, charge transfer in condensed media, flow and transport in porous media, and percolation theory. This chapter is not a review of these areas, so that the reader will be simply referred to the corresponding monographs and reviews at appropriate places in the text.

The key objective of modeling is however straightforward. It should establish the relationship between local distributions of material properties (structure, composition, morphology) and local operational characteristics (reactant concentrations,

pressures, electrostatic potential, reaction rates) with simultaneous evaluation of the global performance of the cell (impedance response, water management, gas and liquid flows, current–voltage relations, stability issues). Experiment can probe these issues separately. Theory can establish links between different disciplines in fuel cell research and integrate information from distinct experimental sources, cf. Fig. 1 (b). A recent feature article in *The Fuel Cell Review* highlights the versatile role of theory in the optimization of fuel cells [1].

There are basically two approaches to fuel cell modeling. One may be conventionally called the *engineering approach*, in which the model includes all imaginable processes which occur in the cell, described by a system of equations, as many as necessary, and with up to a hundred parameters involved. If the majority of these parameters are independently measured, the approach certainly has its own value. However, as explorers first cut narrow pathways in a rain forest, and only then the motorways are built by construction companies, the ''engineering'' approach may not be the most efficient at the first stage of understanding how the cell operates. An alternative approach using detailed physical and chemical theory has some 20 years' history, [2–10] with some 50 years' prehistory of the theory of gas diffusion electrodes (GDE) [11–15]. However, microscopic understanding of the processes in the cell components is still incomplete, the functional maps of cells and stacks are just appearing, the current–voltage relationships at an elementary fragment of the two- or three-phase boundary in the porous composite catalyst layer remain controversial, and so on. Therefore, a simplified, ''physical'' approach to the problem is still valuable and not only for tutorial purposes.

This approach deliberately treats only the most important processes in the cell. It does it in such a way that further fine effects (or at least those that look secondary now, but may appear to bring new features to performance at further analysis) can be later put on top of the ''initial construction''. The approach is thus reduced to *a minimum* number of equations and parameters. With no intention of giving at this point a complete list of references, we refer to several articles that have introduced this approach to PEFC/DMFC [2–10, 16–19].

This chapter reviews the modern status of physical models and their possible applications. Needless to say, the predictions and suggestions based on simplified models must be verified experimentally and the key parameters used in the model should be, if possible, independently measured. This is not always available yet. One of the goals of this chapter is, therefore, to draw attention of experimentalists and engineers to the results of physical models.

Illustrations, for example functional maps in the cell or current–voltage performance plots, shown in this chapter, rest on certain characteristic values of parameters and operation conditions. It is explained how these plots were obtained with reference to original publications, and thus the same maps and graphs may be plotted for another set of input parameters and conditions, which might seem more realistic or more interesting to the reader. Whether the physical models do pave the right way to more complete engineering models can be proved only by a variety of systematic experiments. Thus, the aim of this chapter is to trigger synergetic activities in this direction.

We start with a physical theory of proton transport in PEMs as a function of

water content. This part comprises microscopic aspects of understanding single proton transfer events in aqueous environments, as well as aspects related to the properties of one single pore and the statistical geometry of the pore network in the sample. Based on these fundamental aspects, alternative models of electroosmotic effect and water management in PEMs under operation conditions will be discussed and compared. The models differ in the description of water transport inside the membrane. Diffusion-type models perceive the membrane as a solution of water in a polymer host with the local membrane state being determined by the water activity. On the contrary, hydraulic permeation type models focus on the porous structure aspects, where the major modes of proton and water transport take place in a bulk-water-like fashion inside the pore network.

In the second part, theories of catalyst layers are presented. They account for the transport of feed gas, protons and electrons and rationalize the distribution of the rate of electrochemical reaction. A theoretically derived ''phase diagram'' suggests an optimum thickness of the layer, subject to the basic parameters and the target current regime.

Under certain approximations, using the concepts of percolation theory, the basic parameters can be related to the volume portions of the components of the layer. This offers a relationship between the structure of the porous composite catalyst layer and its performance. An optimum composition (in terms of volume fractions of electrolyte material, carbon and carbon-supported catalyst, and pore space) is a ''Holy Grail'' here. Albeit this goal can still be far away in view of the simplified character of the models used, these models give at least some rational scheme for the understanding of the composition dependence.

Constructed on the same principles, a theory of the complex impedance reveals the interplay between the structure of the layer and its performance. The results demonstrate how the impedance spectra could help to determine the catalyst layer parameters.

In the third part we describe a quasi-three-dimensional (Q3D) physical model of the fuel cell. The model results in the distributions of concentration of gases, currents and reaction rates in each cell cross section, perpendicular to a long meander-like channel, that is, it provides a functional map of the cell.

The fuel cell is basically a two-scale system. The small and large scales are determined, respectively, by membrane-electrode assembly (MEA) thickness and by the length of the feed channel. The Q3D model is designed to investigate the interplay of small- and large-scale processes in PEFC/DMFC, so that the fully 3D model of the cell is split into a model of a cell cross section (internal model) and a model of the flow in the channel (channel model). The two models are coupled via the local current density along the channel and the overall Q3D solution is obtained by iterations.

We discuss the basic equations of the Q3D model and show the results of simulations for the gas-feed DMFC and hydrogen–oxygen PEFC. Then, we present a simplified theory of ''along-the-channel'' feed gas consumption, verified by the simulations and first experiments.

The chapter is concluded with a brief discussion of the guidelines for an optimal PEFC/DMFC design, based on the functional map of the cell and the established laws of feed consumption.

8.2.2 The Membrane

8.2.2.1 Structure and Function of the Membrane

Currently, the best proton-conducting materials contain water as the vital component, utilizing the advantages of the anomalous proton mobility in hydrogen-bonded water networks [20–22]. PEMs are such materials, used as the key components in PEFC and DMFC [23, 24]. Their function is to provide a stable separation of gaseous reactants on the anode and cathode sides and good proton conductivity. State-of-the-art PEMs fulfill these demands due to a bifunctional composition: a solid matrix of polymer backbones is responsible for stability; a network of water-containing pores, emerging in this matrix, provides pathways for protons. There are practically no, or very few, ultrasmall pores available to gas permeation, since pores open up mainly upon the membrane water uptake and swelling that are driven by hydration of hydrophilic residues. The aqueous pathways for protons lead to the unfavorable side effect of molecular diffusion of water or any other small molecules dissolvable in water, such as, for example, methanol.

The effect of polymer morphology on membrane structure and conductance has been shown recently. In Ref. 25 hydrogen-based graft-copolymer membranes were compared in terms of morphology and performance to random copolymer membranes with the same ion content. For the hydrated grafted membranes TEM micrographs revealed a picture of a continuous phase-separated network of water-filled channels with diameters of ~5 nm. In contrast to that, the random copolymer membranes exhibit a reduced tendency toward microphase separation; water is dispersed more randomly within the polymer domains. The water uptake in the grafted membranes is limited to smaller values, since the hydrophobic nature of channel walls impedes the penetration of water into the polymer domains. Nevertheless, their ionic conductivity is of an order of magnitude larger than that of the random copolymer membranes.

An adequate structure of polymer molecules promotes the advantageous phase separation into hydrophobic and hydrophilic domains upon water uptake. The most notable class of membranes based on this principle are the perfluorosulfonic acid ionomers (PFSI), Nafion® [26] and similar membranes [27]. In these membranes, perfluorosulfonate side chains, terminated with hydrophilic $-SO_3H$ groups, are attached to a hydrophobic fluorocarbon backbone. The tendency of ionic groups to aggregate into ion clusters due to the amphiphilic nature of the ionomer leads to the formation of basic aqueous units. At sufficient humidity these units first get connected by narrow channels and then may even fuse to provide continuous aqueous pathways [28].

Exposed to water, the strong polymeric acids dissociate into anionic groups fixed at the side chains and free mobile protons in the aqueous solution. Thereby, the degree of hydration determines the availability and mobility of protons in the water-containing domains.

Figure 2 sketches different length scales on which water of hydration affects membrane structure and transport properties. The bottom picture represents the smallest unit of a membrane that still contains the essential ingredients of the membrane interior: clusters of acidic functional groups and water molecules. The electronic structure of acidic groups and polymer side chains determines the kinetics

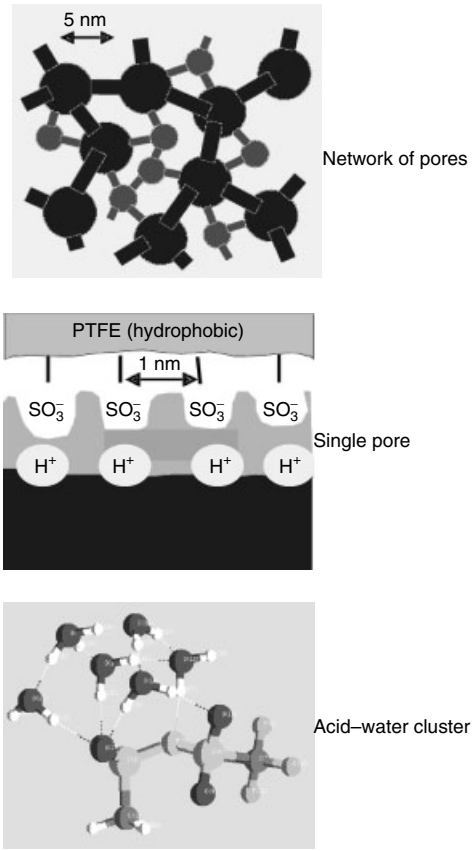


Fig. 2 Sketching different levels on which water of hydration affects the membrane performance. On the molecular level (bottom) explicit acid/water interactions determine dissociation and hydration of acidic functional units. The central picture illustrates the effects of the water structure within a single pore and the arrangement of anionic groups on the pore surface. On the largest length scale, for the network of pores (top), pore size distribution, and connectivity among pores establish the link between water distribution and membrane conductivity.

and energetics of acid dissociation and the structure of water of hydration in the immediate vicinity of the acid groups [29, 30]. These aspects are crucial for membrane operation at low water contents [31, 32]. Electronic structure calculations, for example, density functional theory (DFT) and ab initio molecular dynamics (AIMD), are thus viable tools to validate whether an acidic group would represent an appropriate functional unit in fuel cell polymers. For a recent review see Ref. 33.

Due to computational requirements high-level quantum mechanical calculations are, however, limited to small systems (~10–100 atoms), simple periodic structures and short timescales (~10–100 ps). Larger systems can be studied by empirical molecular dynamics (MD) simulations. Such simulations based on Walbran–Kornyshev model [34] of proton transport in water allowed to study the proton mobility in pores, modelling the interior of well-humidified Nafion, see also Refs. 35–38.

The picture in the central part of Figure 2 demonstrates the significance of collective effects due to the array of anionic groups at the interface between hydrophobic polymer aggregates and water-filled domains. Moreover, it suggests the distinction between different water environments within water-filled pores. In the central region of a pore, proton mobility is anticipated to be similar to the mobility in bulklike water [33, 39–42]. Near the pore surface a strong influence on water structure and proton mobility due to the density and arrangement of anionic groups is expected [34, 35, 37, 38, 43, 44]. For instance, it was found that the more flexible are the side chains, that is, the more mobile are the ions of side chains although anchored to the corresponding points of the backbones, the faster is the proton transport.

There exists a trade-off between the expense at which protons are available and the ease with which they move in different pore regions. Proton concentration decreases from the surface to the center, but their mobility is highest in the pore center and low in the proximity of anionic groups. The distributions of proton density and mobility within the pore determine the dependence of proton conductivity on pore radius. This dependence was studied in detail in Refs. 40, 43. Increase of the pore radius shifts the balance from surface-type to bulk-type conductivity.

On the largest length scale, top picture of Fig. 2, the distribution of water in the membrane is depicted as a porous network. The latter is characterized by a pore size distribution (psd) and a tortuousity factor ξ, which accounts for multiple interconnectivity and bending of pathways in the network. The distribution of pore radii translates into a distribution of pore conductivities. Via this correspondence, the distribution of water in the membrane finally determines its transport properties, namely, proton conductivity and water diffusivity.

Since PEM, like living organisms, need water to function and the amount and state of water are critical for an efficient operation, secondary requirements on this type of fuel cell membranes emerge. These include the necessity of sufficient humidification and the ability to retain water under operation conditions. Associated problems under fuel cell operation include the electroosmotic transport of water to the cathode side accompanied by dehydration at the anode side [45]. In the cathode the accumulation of water at high current densities, typically greater than 1 A cm^{-2}, causes performances losses due to blocking of catalytically active sites and restriction of oxygen transport.

8.2.2.2 Membrane Conductance

Several models of the random pore space and its evolution upon membrane swelling are suggested in Refs. 39, 40 to bridge the gap between the conductivity of a single pore and the conductance of a membrane. Combined with experimental data on psds they reveal the features that cause good or bad membrane performance.

It was possible to rationalize the family of Arrhenius plots measured for Nafion® 117 at different water contents [46]. Under an assumption that the surface conductivity has higher activation energy, supported by microscopic considerations in Refs. 40, 43, the Arrhenius slope should become steeper with the decreasing amount of water in the membrane [39], that is, the smaller the amount of the ''bulk'' water that we have in pores. Activation energies obtained from these plots are $\approx$0.1 eV for the largest possible water contents (Activation energies of proton transfer in water, estimated from nuclear magnetic resonance relaxation times, are $\approx$0.1 eV [47].) and 0.3–0.4 eV at small water contents. How to rationalize this variation?

8.2.2.2.1 Proton Transfer in a Single Pore

The kinetic parameters of the elementary act of proton transfer in a pore were studied for the model system of a thin electrolyte slab between two infinitely thick polymer walls in Ref. 43. The fixed anionic charges (SO_3^- groups) were represented as a regular square lattice array of point charges at the polymer/water interface. For this configuration, electrostatic contributions to reorganization energy, E_r, and work terms (reaction Gibbs free energy), ΔG, of the elementary act were calculated. Within the charge-transfer theory, the activation energy, G_a, of proton transfer may be evaluated by the Pekar–Marcus

equation [48],

$$G_a = \frac{(E_r + \Delta G)^2}{4E_r} \qquad (1)$$

(In principle, Eq. (1) is only valid in the diabatic regime. For proton transfer, that is considered to occur in a partially adiabatic regime, it is not strictly accurate. Nevertheless, it can still be expected to give a good upper estimate of the activation energy.)

The water reorganization energy is mainly determined by the size of the water clusters between which the proton is transferred (H_3O^+, $H_9O_4^+$) and the effective distance over which the proton is shifted during the process. The contribution of the medium reorganization outside the reactant water clusters (outer-sphere contribution) is calculated in Ref. 43, 44 as a function of reactant cluster sizes, transfer distance, pore size, and dielectric properties of the contacting media (water, polymer).

Assuming $H_9O_4^+$-reactant complexes with radius 3 Å, a transfer distance 3 Å (overlapping reactant clusters), high frequency and static dielectric constants of confined water $\varepsilon_{w,hf} = 5.5$ and $\varepsilon_{w,st} = 20$, the outer-sphere contribution to the reorganization energy was obtained as $E_r^{os} = 0.16$ eV. The pure geometrical effects due to finite pore size and proximity of the reactant clusters to one of the interfaces were found to give only negligible corrections to E_r^{os}. The confinement affects the properties of pore water, by freezing out its orientational degrees of freedom, thereby reducing the effective dielectric constant of the pore interior. With a static dielectric constant $\varepsilon_{w,st} = 8$, $E_r^{os} = 0.07$ eV is obtained in Ref. 43, 44. The confinement of pore space can thus be expected to lead to an effective reduction of the outer-sphere reorganization energy. Attempts to estimate the reorganization energy contribution due to the local modes inside the reactant spheres on the basis of molecular-level calculations, using ab initio or pseudopotential based MDs, have also been reported [36, 49, 50].

For proton transfer between symmetric bulk water like states, the reorganization energy would be the only contribution to the activation energy. Calculated activation energies for bulklike proton transfer (≈0.1 eV) are in good agreement with experimental values for Nafion® 117 at high values of water content, suggesting that in this regime the majority of protons acquire the bulklike mobility.

The reason for the higher activation energy of surface conductivity could first of all be the strong coulombic work term ΔG for proton motion along the array of SO_3^- groups. This contribution is essential, if the proton is transferred from an SO_3^- group to an intermediate water molecule in the rate-limiting step.

Dependencies of ΔG on the average separation l between SO_3^- groups and on the distance a between the plane of proton transfer and the negatively charged interface were studied theoretically in Ref. 43, 44. Using an approach based on the Poisson–Boltzmann equation, modulations of the electrostatic potential and of the distribution of mobile protons in the proximity of the charged anionic sites were calculated.

The results revealed a strong dependence of ΔG on a. ΔG, which is determined by the modulation of electrostatic potential due to the array of fixed negative charges on the surface, becomes small (<0.05 eV) at $a > 3$–4 Å. Moreover, a higher surface density of fixed negative charges, that is, smaller l, results in shallower potential wells for the protons in the proximity of these groups. At $l \approx 15$ Å

the work term due to electrostatic potential modulation is ≈0.5 eV (at a distance ≈1 Å, corresponding to the first monolayer of water on the interface), whereas at $l \approx 5$ Å this work term is only ≈0.05 eV.

Note that calculations in which SO_3^- groups are considered as fixed point charges overrate the role of ΔG. A modified Poisson–Boltzmann–Ansatz in Refs. 37, 38 takes into account finite size form factors due to more realistic charge distributions for the sulfonate groups and thermal fluctuations of the group positions. Due to these effects, the ''coulombic profile'' along the pore surface becomes considerably flatter and, thus, ΔG is smaller. Additional contribution to the activation energy in narrow pores, may come, however, from the increase of the frequency of local modes contributing to reorganization energy [51]. The variation of the activation energy for proton transport with pore width is consequently found to be smaller, which is consistent with findings from MD simulations [34, 35, 37, 38].

So far, the picture of proton transfer in water-filled pores at high water content looks consistent. In the pore center, only water reorganization energy contributes to the activation energy of proton transfer, which occurs between symmetric states. This reorganization energy is mainly determined by the size of the water clusters between which the proton is transferred (H_3O^+, $H_9O_4^+$) and the effective distance of the charge transfer. For proton transport in the surface layer (≈3 Å thick) the work term ΔG has to be taken into account in Eq. (1), leading to the higher observed values of G_a. ΔG exhibits a strong dependence on the separation between SO_3^- groups but is strongly diminished by thermal fluctuations of their positions: the more labile structural arrangements of SO_3^- groups will provide lower activation energies of proton transfer.

Alternatively the pathway of surface conductance under conditions of very low water contents, where the formation of a native hydrogen-bonded network is impeded, may proceed via a restricted mobility of SO_3^- groups. Fluctuations of these groups toward each other can create a transition state in which the proton is shared between the two groups and eventually is transferred before the break-up of the transition state. In this case, proton transfer takes place between symmetric states, and no work terms are involved [52]. There is, however, a price for the formation of the transition state. It costs to bring the ends of the side chains with SO_3^- groups on them closer together, which may be particularly hard in narrow pores with crystal-hydrate-like water structure inside. Recently performed AIMDs simulations on a trifluoromethane sulfonic acid monohydrate solid, which represents a viable reference for minimally hydrated Nafion® with restricted mobility of SO_3^- groups, suggested that the activation energy of this transition state is ≈0.3–0.4 eV[53].

8.2.2.2.2 **Proton Conductance as a Function of Pore Radius** A basic model of proton transfer in membrane pores must distinguish a surface mobility with activation energy ≈0.4 eV, exhibiting a strong dependence on area density and mobility of the surface charges, and a bulk water like mobility with an activation energy ≈0.1 eV. Based on these assumptions, the proton conductance of a pore was calculated in Ref. 40 as a function of the pore radius. The radial distribution of protons in the pore was obtained in a closed form expression from a mean field approach. The latter can, of course, give an error, if the average radius of a volume per proton

in the pore is larger than the Debye length in the atmosphere of free protons, which is typically the case. But it still could be used for a crude estimate. This distribution has higher density near the negatively charged surface and lower density in the middle of the pore. Increase of the pore radius shifts the balance from surface-type to bulk-type conductivity, with a simultaneous decrease of the effective activation energy.

The main effect that leads to the calculated increase of pore conductance with radius is the redistribution of protons upon swelling toward higher concentration in the bulk. An estimate in Ref. 40 has demonstrated that at pore radii of ≈30 Å about 75% of the protons move in the bulk. Furthermore, in accordance with small angle scattering data (SAXS, SANS) [54] it was taken into account that, as a rule, the mean separation of SO_3^- groups increases upon the expansion of a pore. Together these two effects promote the crossover from surface conductance at smallest radii (<5 Å) to bulk conductance at large radii. These results are in line with the experimentally detected decrease of activation energy upon pore swelling [46].

8.2.2.2.3 Effects of Membrane Morphology

There is still a long way between the conductance of a single-pore and the macroscopic membrane conductivity. Water fractions in the membrane form tortuous pathways with ramified boundaries. Upon water uptake the system continuously evolves by the swelling of individual pores and the creation of new connections between pores. For modeling purposes, the highly interconnected porous structure may be subdivided into elementary segments resembling lamellae, cylinders, or spheres.

It would have been natural to assume that cylindrical pores prevail in the swollen membrane. Upon dehydration, the number of pore constrictions increases, possibly leaving structures that look like networks of spherical pores (inverted micelles), connected by narrow cylindrical channels. The latter type of structure was proposed earlier by Gierke et al. [54] on the basis of small angle scattering data (SAXS, SANS). Small angle scattering data of Gebel et al. on the structural evolution of water swollen membranes supported the inverted micellar structure model [55]. At very large water contents (larger than 50% by volume) a structure inversion to a connected network of polymer-rod-like particles is proposed in Ref. 55. Very recent SANS and neutron diffraction measurements performed for Nafion® and PEEK type membranes revealed a variety of porous structure patterns strongly dependent on the membrane prehistory [56]. For a number of membranes at maximal saturation a network including clusters of overlapping spheres rather than a more homogeneous network of cylindrical channels has been revealed.

For Nafion®, decisive new experimental insight has been obtained by small angle X-ray scattering, suggesting a modified structural model [57, 58]. Whereas previous models focused on the evolution of water-containing pathways within an inert and structureless polymer host, the new data indicate that polymer rods and their aggregation should be considered explicitly as the structure forming elements. This view opens new intriguing perspectives for predictive theoretical modeling of the formation of random heterogeneous membrane structures, based on the consistent treatment of interactions between polymer components.

In view of the vague experimental situation regarding pore-space morphology, several models of the pore space and

its evolution upon water uptake were addressed in theory [39, 40]. These are the: (1) simplistic model of a monodisperse sample (consisting of parallel cylindrical pores with uniform cross sections), (2) model of uniform parallel pores with varying radii, (3) serial model with sections in sequence having randomly distributed radii, (4) random network model. For the random network models effective medium theory provides approximate expressions of the conductivity. Combined with available data on pore-radius distributions, these models reveal the structural features that cause good or bad performance.

Moreover, different approaches to the pore-space evolution were considered. One approach was based on a continuous swelling process of the network [40]. Starting from a network of cylindrical pores, elementary segments of the pores are subject to an iterative, random swelling procedure. The swelling is controlled by a single parameter that determines the skewness of the resulting psd. In the continuum limit this swelling procedure results in psds with log-normal shape, which are viable psds in ultrafiltration membranes [59]. As will be discussed in Sect. 8.2.2.7.4, experimental data for psds of Nafion® 117, plotted in Fig. 5, are approximated well by a log-normal distribution [60].

Implying certain restrictions on the previously described swelling process leads from the continuous swelling approach to the two-state model that was considered in Ref. 39. In this approach, each pore is allowed to swell only once: from the small dry state (only residual water) to the large water-filled state, which has an equilibrium radius determined by polymer–water interactions. A phenomenological model of swelling (including the merging of small pores into large ones), governs the humidity dependent balance between the populations of the two states.

The theoretical approaches to the pore-space evolution upon water uptake permit a straightforward incorporation of the dependence of single-pore conductance on radius. Calculations within each of the models result in an effective single-pore conductance, which is proportional to the specific membrane conductivity. Therefore, closed form relationships between porous structure, water uptake and membrane conductivity have been calculated [39, 40]. The highest conductivities are obtained within the simplistic model. Between the two main model variants, that is, the serial model and the random network model, better performances are found with the latter. At given water content, narrow psds result in better performances than broad polydisperse psds. In the latter case the evolving membrane structure becomes rather heterogeneous upon swelling, whereas narrow psds resemble the simplistic model with homogeneous water distribution.

Numerical estimates of proton conductivity on the basis of this theory reproduce the main experimental trends. In Ref. 39 it was found that the two-state model with the account of swelling properties gives results that are in agreement with experimental data for Nafion® 117.

Differences in conductivity data of modified perfluorosulfonate membranes can be related to structural differences on the basis of the pore structure models. Smaller equivalent weights (e.g., for Nafion® 105, Dow, Membrane C), that is, higher specific ion content, lead to superior performance compared to Nafion® 117 due to narrower psds and, thus, more homogeneous water distributions.

In its present form the theory and simulations of proton transport in single

pore can give only limited indications on what should be the modification of the polymer for a superior performance at all water contents. The recipes are simple: make side changes more flexible and denser, and you get livelier proton transport. Recent findings of P. Commer (Dissertation, Research Center, Jülich) show that simultaneously the diffusion of methanol is impeded by such modifications which is good news in the context of DMFC applications. More generally, this theory gives a useful framework for a better understanding of the main factors that influence membrane conductivity in a pore as a constitutive element of the proton-conducting network in membranes. This framework can be used for treatment of experimental data and for the formulation of the tasks for computer simulations. All this is valid as far as we have pores in the membrane. But do we really have them?

8.2.2.3 **Pores or Necks?**

Recently, Ioselevich, Kornyshev, and Steinke (IKS) suggested a more detailed morphological model of Nafion-type ionomers [61], which has questioned the existence of pores in homogenous ionomers such as Nafion at least in ''normally'' hydrated state. The model was based on neutron scattering data showing an approximately tetrahedral coordination of water droplets – the cores of inverted micelles – with distances between their centers detected from the Bragg peaks [55]. The IKS model was in essence a development of Gierke's model, albeit with inverted micelles not free but coupled to the geometry of the polymer backbones. It begins with the question: How can one construct an array of inverted micelles built by hydrated side chains, if the latter are attached to the backbone chains arranged in bundles? Which kind of structures build up when the bundle persistence length is considerably larger than the size of the micelles? A quasicrystalline model of this arrangement has been developed, with probable unit cells as depicted in Fig. 3.

Two possible types of cages exist, in which the strings represent the backbone chains or their bundles. The cages are capable of providing a short-range-ordered system of four coordinated, inverted micelles composed of hydrated side chains (not shown) pointing from the strings toward the interior of the cages. These micelles keep water droplets (not shown) with protons of dissociation encapsulated inside the cage. The channels, when they form, bridge water droplets, through the windows in the cages and are, in this model, nothing else than narrow aqueous necks connecting the droplets.

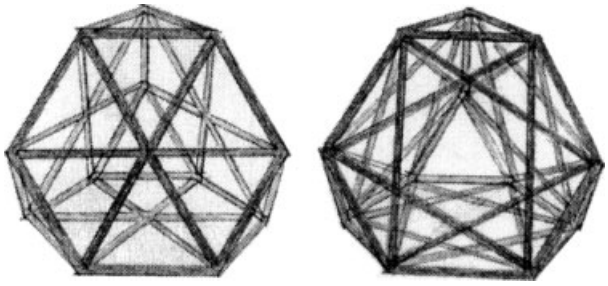


Fig. 3 Possible unit cells of the Ioselevich–Kornyshev–Steinke quasicrystalline model of perfluorinated ionomers.

This model managed to rationalize the observed correlation between the macroscopic swelling of the membrane volume, V, and the expansion of the average distance d between the micelles: $V \propto d^2$. Channels that could be built through the windows of the cages will be very narrow. Upon membrane swelling, the cages expand by strings sliding along each other. At the beginning, the size of the droplets grows, but each droplet is still encapsulated within each cage, disconnected from other droplets. With further water uptake, the droplet ejects water into the windows of the cages, building small cathenoids (minimal surface forms) adjoining the neighboring droplets. The analysis shows that this process will likely take place as a first-order transition [62]. At the transition, the system may shrink slightly as some water from the droplets will be taken to build channels. With further water uptake, the system will swell again, and both the droplets and cathenoids will increase in size continuously. A theory of this phenomenon is currently under development.

So how can one use this knowledge practically? This is still not obvious, though it does seem that we at least have a better understanding of the nature of the proton-mobility dependence on the water content. To sum up: if the channels evolve in the beginning as extremely narrow units (less than 0.5-nm radius for the narrowest part of the cathenoid) and remain narrow even in the ''mature state'', it is clear why the activation energy of the proton mobility (which is entirely controlled by the necks) depends, dramatically, on the water content. And the more flexible the side chains, the higher the proton mobility, since fluctuations of the chains will support the necks, reducing their surface tension; there could also be proton transport promoted by side-chain fluctuations.

Finally the question arises, are there any channels or pores apart from the necks between the droplets? The IKS model is probably good for dense ionomers with no, or small number of, defects (with defect density lying below the percolation threshold), so that the bottleneck of the proton transport is the proton transfer along the necks. It naturally explains the dramatic variation of activation energy of proton mobility with water uptake, which is difficult to do, if the *size of the droplet* is taken as the *characteristic size of the channel*, the quantity considerably larger than the size of the neck. In reality, the quasicrystalline order may be distorted or interrupted by disordered domains. It is often argued that most of the ionomers are, like glasses, materials far from equilibrium and their properties often depend on ''equilibration prehistory''. In such a system one would expect a hierarchy of scales rather then one rigorous quasiperiodical motif. Bragg peaks of diffraction data seem, however, to be more in harmony with the quasicrystalline model, although this case is still far from being closed.

The IKS model is more natural for the low water content, with backbones not ''plasticized'' by excess water, whereas in water-saturated membranes the quasicrystallinity may not be the dominating motif at all. The main assumption of the quasicrystalline model is that the persistence length of the bundles of the backbone chains is at least several times greater than the distance between neighboring micelles. This condition may break down with the increase of water content. The IKS model helps to understand the ''birth'' of conductivity and its evolution with increasing water content, but a more complete theory of this evolution should invoke the

effects responsible for the ''dissolution'' of the quasicrystalline order. This dissolution probably goes via merging of the neighboring micelles into larger and preferentially cylindrical blocks. Such a theory is difficult to build on a microscopic level. On a phenomenological level, a similar type of evolution of aqueous pore space had been considered in Ref. 39.

8.2.2.4 Membrane Development

Nafion® and similar conventionally used separator membranes in PEFCs offer excellent performance characteristics, but they suffer from inherently high production costs. Due to the critical dependence of membrane performance on sufficient hydration, their conductivity starts to deteriorate dramatically at temperatures $>100\,^{\circ}C$ [24]. Moreover, fuel crossover and anode poisoning lead to considerable performance losses of the fuel cell if fuels other than pure hydrogen are used [63]. One of the major targets in fuel cell research for the future is, therefore, the development of high-temperature membranes, which are free of liquid water [24, 64, 65]. In fact, stable membrane operation at elevated temperatures in the range of 110 to $150\,^{\circ}C$ would be beneficial in several respects, facilitating heat management for the overall system, increasing CO tolerance of the anode, enhancing reaction rates in both electrodes and, thus, reducing fuel crossover. Membrane modifications are possible on all the different levels in Fig. 2.

The main efforts are directed toward better performing, more cost-effective membranes, based on (1) reduction of membrane thickness [66, 67], (2) control of composition and swelling properties via the equivalent weight (or ion-exchange capacity) [67] or via reinforcing the nonconductive matrix by means of cross-linking, (3) modifications of the hydrophobic backbone, for example, using nonfluorinated or partially fluorinated, radiation-grafted polymers [68–70], or embedding the acidic functional units into a stable inorganic container [65], (4) using polybenzimidazole (PBI) films doped with phosphoric acid, which exhibit good proton conductivity ($\sim$0.01–0.04S cm^{-1}), very small electroosmotic drag coefficient, low methanol crossover, excellent chemical and thermal stability, and good mechanical flexibility at elevated temperatures ($150–200\,^{\circ}C$) and low humidification [71], (5) solid inorganic acid compounds such as $CsHSO_4$, which offer the advantages of anhydrous proton transport and high-temperature stability (their viability for fuel cell applications has still to be demonstrated) [72], (6) replacement of liquid water by immobilized proton solvents such as imidazole, which generate similar proton mobility as hydrated perfluorosulfonic polymers, but potentially reduce the problems due to mass flow (fuel crossover, electroosmosis) [73], (7) variation of the functional acidic group and associated modifications of the properties of acid dissociation and hydration [30, 74].

Novel approaches in PEM synthesis focusing on cheaper, usually fluorine-free PEMs and membranes capable of sustained fuel cell operation at elevated temperatures, have been reviewed in Ref. 75. Mature fuel cell membranes have been casted from sulfonated poly(arylene ether sulfone) (BPSH) random (statistical) copolymers [76].

They exhibit excellent thermal and mechanical properties, enhanced chemical stability and conductivities similar to Nafion® ($\sim$0.08 S cm^{-1} at $30\,^{\circ}C$). Blending of these polymers with heteropolyacids

facilitates controlling membrane morphology and water retention. Conductivities of $\sim$0.17 S cm^{-1} can be obtained. Moreover, a novel class of graft-copolymer membranes (PS-g-PSSA) with controlled polymer structure and water content has been studied [77]. Effects of density and length of graft chains on membrane morphology, water uptake and conductivity have been rationalized. Conductivities of $\sim$0.24 S cm^{-1} have been realized. Although the latter PEMs are not sufficiently stable at elevated temperatures, they are highly insightful for the structural PEM design.

In practice, trade-offs between optimization of different membrane functions have to be accepted. For instance, the immobilization of the proton solvent will impede the leaking out of solvent and, thus, help to avoid membrane dehydration and cathode flooding. On the other hand this may only be achievable at the cost of lower proton conductivity. A good theoretical understanding of mechanisms of proton mobility in various aqueous and nonaqueous environments is thus of vital importance.

8.2.2.5 Water Transport and Electroosmosis

In PEM fuel cells, fed with neat hydrogen on the anode side, water fluxes in the membrane play a crucial role for the overall water balance of the cell. For the following considerations of this problem it will be supposed that the essential ex situ membrane properties are known. Based on this knowledge the membrane performance in the fuel cell will be studied. Effects of membrane properties and externally controlled conditions will be rationalized. This kind of understanding of structure versus function provides diagnostic tools to check the suitability of membranes for fuel cell applications and yields criteria for properties that a good membrane should have.

Ideally, the membrane should be filled homogeneously with water at the value of water content that gives the best proton conductivity. In particular, models dwelling on fuel cell operation usually consider the membrane a homogeneously hydrated medium with constant transport coefficients [2, 3]. This approximation corresponds to the assumption of an ''ultrathin'' membrane [5, 6].

However, in the cell the membrane hydration is affected by generic fuel cell processes, including the supply of humidified reactant gases to the electrodes, electroosmotic drag of water from anode to cathode, backtransport of water in the membrane, and production of water at the cathode. It is, therefore, generally important to consider the internal membrane water balance self-consistently and relate it to the membrane microstructure.

Stationary fuel cell operation requires a steady flow of protons through all membrane cross sections, perpendicular to the transport direction. Proton flow induces water transport from anode to cathode by electroosmotic drag [78]. Taken alone, this effect would lead immediately to membrane dehydration and to a drastic increase of its ohmic resistance. However, accumulation of water on one side of the membrane inevitably causes a backflow of water. The balance between this backflow and the electroosmotic flow leads to a stationary profile of water across the membrane.

Once the water depletion becomes so strong upon increasing current passage that right at the anode the water content drops down to the ''dry membrane level'', the overall resistance of the membrane becomes very high, and practically no further growth of the current density is possible.

8.2.2.6 Modeling Approaches to Water Management

There are different approaches that incorporate the water balance in the membrane into models of fuel cell performance. They rest on different concepts of membrane microstructure. As a common feature they use local values of transport parameters which are functions of the local water content, w (volume fraction of water relative to the total membrane volume).

The existing models can be grouped in two principal categories, ''black box'' models and ''structural'' models. Within the empirical ''black box'' models the membrane is considered a continuous, nonporous phase in which water of hydration is dissolved. An effective diffusion coefficient which is a characteristic function of the water content controls the water flux.

The majority of recently published papers are based on this kind of effective diffusion model [7, 10]. In the models of Springer et al. [7] and later Nguyen et al. [10], the local membrane state is determined by the local activity of water which is in thermodynamic equilibrium with surrounding water vapor. Diffusion of water driven by the activity gradient balances the electroosmotic flow. Under stationary conditions this results in a characteristic profile of w across the membrane, with lowest values in the proximity of the anode.

Fuller and Newman [4] based their model of coupled proton and water transport in PEFC on the theory of concentrated solutions, wherein the effective diffusion constant was determined from the hydraulic permeability. Based on flux measurements of Fales and Vanderborgh [79], the model of Fuller and Newman used, practically, a w-independent permeability. Therefore, their results are rather similar to those of Springer et al. [7].

Structural models emerge from the notion of membrane as a heterogenous porous medium characterized by a radius distribution of water-filled pores. This structural concept of a water-filled network embedded in the polymer host has already formed the basis for the discussion of proton conductivity mechanisms in previous sections. Its foundations have been discussed in Sect. 8.2.2.1. Clearly, this concept promotes hydraulic permeation (D'Arcy flow [80]) as a vital mechanism of water transport, in addition to diffusion. Since larger water contents result in an increased number of pores used for water transport and in larger mean radii of these pores, corresponding D'Arcy coefficients are expected to exhibit strong dependencies on w.

Refs. 6, 81, and 82 report that D'Arcy flow of water was observed under hydraulic pressure gradients across ion-exchange membranes. D'Arcy coefficients were measured. In the context of water management in fuel cells, the balance between electroosmotic and D'Arcy flows was theoretically considered in Ref. 6. However, the water management analysis was performed under the assumption of a saturated membrane, that is, no variation of w was taken into account.

Following the idea of Yu. Volfkovich, a model of stationary water flows in the membrane with account of porous structure-related aspects and inhomogeneous water distribution was developed [16, 83]. This model will be presented in some detail below. Its implications on water-content profiles and current–voltage performance under fuel cell operation conditions will be compared to the effective diffusion models.

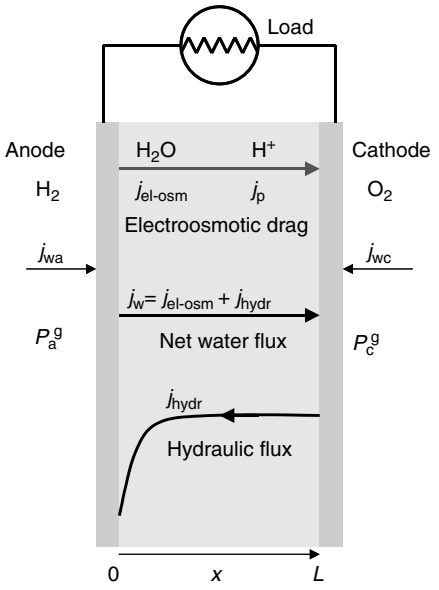


Fig. 4 Membrane in operating fuel cell: a scheme depicting the transport processes, that is, gas supply, proton current, modes of water flow.

Note that diffusion models and hydraulic permeation models have their own caveats: the membrane is neither a homogeneous acid solution, nor is it the well-structured porous rock. Critical comparison of the results of the two approaches with each other and with experiments, is of crucial importance for understanding the membrane functioning within the cell and developing the strategies on water management and optimized membrane properties.

A nonconventional view of membrane microstructure, which neither conforms with the solution nor with the porous rock picture, was recently suggested in Ref. 84. Classical MDs simulations on microstructure and molecular mobility in swollen Nafion® membranes revealed a picture of a rather dynamic structure of water clusters with temporary formation and break-up of water bridges between them. The frequency of intercluster bridge formation was found to be consistent with the experimental transport coefficients through the membrane.

8.2.2.7 Hydraulic Permeation Model

Generally, the membrane under operation conditions should be considered inside the fuel cell environment. Works on simulation of water management within the whole cell or stack usually utilize a rather simplified picture of the membrane. The complementary objective of the model described in this section is, however, to reveal primarily the effects of water distribution across the membrane. Therefore, appropriate boundary conditions on the anode and cathode sides of the membrane are used to simulate the membrane performance inside the cell (cf. Fig. 4).

For this purpose the electrochemically active electrode layers at $x = 0$ (anode side) and $x = L$ (cathode side) are considered as infinitely thin. Gas transport in the backing layers is assumed to be facile (high permeability). Inhomogeneous distributions of feed gases along the electrode planes may thereby be ignored. All local properties, thus, depend only on the coordinate x perpendicular to these boundary planes, and the problem is reduced to a one-dimensional one. Moreover, heat transfer phenomena are ignored, assuming that thermostatic conditions are externally provided.

The heterogeneous model of PEMs implies the existence of a water-filled porous network, which reorganizes upon water uptake. This reorganization has two major impacts on transport properties: the increase of single pore cross-sectional areas, available for proton and water transport,

and the evolution of orientation and topology of pores in the network. Therefore, the phenomenological description of proton current and water management relies on the following experimental information: (1) water content w as a function of liquid pressure P^{l} in the pores (or more generally as a function of the free energy of water binding with the membrane material), (2) pore-radius distribution $w(r)$ [60, 81], (3) electrolyte conductivity as a function of water content, $\sigma_{\mathrm{el}}(w)$ [39, 46, 85], (4) electroosmotic drag coefficient, n, that is, the number of water molecules transported through the membrane together with each proton, and its dependence on w [86, 87]. The dependencies (1) and (2) determine mutual relations between water content, membrane morphology, and water transport.

8.2.2.7.1 **Concept of Capillary Pressure** Pore-radius distributions and ab-/desorption isotherms are important structural characteristics of generic porous media [80, 88]. The absorption isotherm provides a relation for the liquid uptake of a porous medium under controlled external conditions, viz., the pressure of an external fluid. Within a bounded system, such as a cylindrical tube, a discontinuity of the pressure field across the interface between two fluid phases exists. The corresponding pressure difference is called *capillary pressure*, P^{c}. In the case of contact between gas phase, P^{g}, and liquid water phase, P^{l}, the capillary pressure is given by

$$P^{\mathrm{c}} = P^{\mathrm{g}} - P^{\mathrm{l}} \tag{2}$$

According to the thermodynamics of capillary condensation, the Kelvin–Laplace equation determines the free energy of water binding to the membrane [80],

$$\Delta G_{\mathrm{w}} = -RT \ln\left(\frac{P}{P_{\mathrm{s}}}\right) = P^{\mathrm{c}} V_{\mathrm{m}} = \frac{A V_{\mathrm{m}} \gamma \cos\theta}{r} \tag{3}$$

where P/P_{s} is the relative humidity (P is the water vapor pressure and P_{s} is the saturated water vapor pressure, corresponding to a flat meniscus), V_{m} the molar volume of water, r the characteristic radius of the meniscus, γ the surface energy of the liquid–vapor interface, θ the wetting angle, and A the geometry parameter of the pore. For a cylindrical pore r is the radius of the tube and $A = 2$. Within a hydrophilic pore P^{l} will be smaller than P^{g}. Increasingly larger pores will be filled with water if P^{l} approaches P^{g}.

Equation (3) establishes a relationship between controlled water vapor pressure P, capillary pressure P^{c} and pore radius r. Equation (2) relates P^{c} to the external gas pressure P^{g} and the internal liquid pressure P^{l}, whose gradient is the driving force of water flux. The presented formalism, thus, provides a closed set of equations that relate the stationary water profiles in the membrane with its porous structure.

These concepts have been routinely employed to determine psds of genuine porous media [88]. A difficulty arises when they are applied to PEMs, since these membranes do not possess an intrinsic porosity. Instead, pores in them are created by the water of hydration, whereas in the dry state the pore network collapses. Gas permeability of PEM is very small. Thus, only with a certain degree of tolerance can one speak about three-phase capillary equilibria, implied in the Laplace equation. It is rather a semiempirical phenomenology, that allows one to relate the liquid pressure (the driving force of the hydraulic permeation)

with the pore radius, itself depending on water content.

At the same time, there is no crude physical controversy in this approach. The largest pores are those which are obtained at given P^c by capillary condensation. Gas filled pores, that is, pores with larger radii than those corresponding to Eq. (3) do not exist, and water-filled pores do not conduct gas. Equilibrium with the gas outside the membrane is established.

Another conceptual difficulty is the question of adequacy of macroscale thermodynamic parameters in nanometer size pores, where the molecular-scale effects prevail. Much work has been done to check the limits of the validity of the Kelvin–Laplace equation [32]. Predictions of this equation have been verified in experiments and MDs simulations in pores with capillary diameters as small as 2 *nm* [89, 90].

A close set of equations was formulated in Ref. 16, related to the capillary pressure isotherms determined by the method of standard porosimetry [60]. In the latter procedure, the equilibrium amount of the wetting liquid is measured in the porous sample under study. Simultaneously, the amount of the wetting liquid is measured in the standard specimen with a genuine porous structure, in which the capillary equilibrium is established. The standards are kept in thermodynamic equilibrium with the sample. The comparison of the amount of wetting liquid in the membrane with the pore-radius distribution in the standards, enables one to record (with a minimum of theoretical assumptions), the volume-size and surface-size distribution curves, specific pore-space surface area, and absorption isotherm in the membrane of interest, for various wetting liquids.

The relative humidity primarily controls the capillary pressure which determines the water filling of the membrane, cf. Eq. (3). Gradients in capillary pressure, related to gradients in pore radii, are the primary internal driving forces of water flux in the membrane, considered in Ref. 16. An additional external gas pressure gradient may be superimposed on these internal gradients.

It is assumed, that the capillary isotherm,

$$w = f(P^c) \tag{4}$$

determined by the method of standard porosimetry, is valid in each representative elementary volume (REV) with a size much smaller than the macroscopic membrane dimensions, but much greater than the size of the largest pore.

Under conditions of current flow w is a function of the lateral coordinate x. Moreover, P^c will also be a function of x, as implied by Eq. (4). This, in conjunction with Eqs. (2) and (3), establishes mutual relations between local values of w_x, P^c_x, P^l_x, and the r_x. For different gas pressures at the anode, P^g_a, and cathode, P^g_c, one may use an approximation with a linear gas pressure profile,

$$P^c_x = -P^l_x + P^g_c - \Delta P^g \left(1 - \frac{x}{L}\right), \qquad \Delta P^g = P^g_c - P^g_a \tag{5}$$

and, therefore,

$$w_x = f\left[-P^l_x + P^g_c - \Delta P^g \left(1 - \frac{x}{L}\right)\right] \tag{6}$$

The external gas pressure gradient will, thus, be superimposed as a driving force of water flux on the internal pressure gradients that develop due to electroosmotic effect and capillary forces. Note that gas pressures in anode and cathode chambers can be controllably varied.

8.2.2.7.2 **Basic Equations** In the following sections, internal mechanisms of water flux in the membrane due to capillary forces will be explored, considering $\Delta P^{g} = 0$. The effect of nonzero ΔP^{g} is presented separately in Sect. 8.2.2.8.2 Two equations determine the distribution of water in the membrane and its effect on current–voltage performance under operation conditions.

The balance equation of water flow is determined by

$$j_w = n(w_x) j_p - F c_w \frac{K(w_x)}{\mu} \frac{dP_x^l}{dx} \qquad (7)$$

Here, the electroosmotic flow is proportional to the proton current density j_p with a drag coefficient $n(w_x)$. D'Arcy flow as the mechanism of water backflow proceeds in the direction of the negative gradient of liquid pressure, which (for $\Delta P^{g} = 0$) is equal to the gradient of capillary pressure. The density of water, c_w, and the viscosity, μ, are assumed to be independent of w. The transport coefficient of D'Arcy flow is the hydraulic permeability $K(w_x)$.

For $K(w)$ a percolation-type modification of the Hagen–Poiseuille–Kozeny equation is considered

$$K(w_x) = \xi \frac{(w_x - w_c)\rho(r_x)}{8} \Theta(w_x - w_c) \qquad (8)$$

where $\xi(<1)$ is the inverse tortuousity factor. For an isotropic tortuousity in three dimensions $\xi = 1/3$. Θ is the Heaviside step function which accounts for the existence of a percolation threshold w_c in the water-permeating network [91].

$$\rho(r) = \frac{1}{w} \int_0^r \frac{dw(r')}{dr'} r'^2 dr' \qquad (9)$$

is the mean square radius of pores that contribute to the water flow in a REV with local water content w.

As a boundary condition to Eq. (7), P^l can be fixed either on anode, P_a^l, or cathode side, P_c^l. In practice, that is, under operation conditions, it is justified to suppose full saturation of the membrane at the cathode side, due to the cathodic water production and the direction of electroosmotic drag to this side. Therefore,

$$P_c^l = P_c^g - f^{-1}(w_s) \qquad (10)$$

where w_s is the saturation water content.

The net water flow is considered to be a parameter under control,

$$j_w = j_{wa} = j_{wc} - \frac{1}{2} j_p \qquad (11)$$

where the last term on the rhs accounts for water production in the cathode. j_{wa} and j_{wc} are controlled water vapor flows through anode and cathode chambers.

The current–voltage relation is obtained from the migration equation,

$$j_p = -\sigma_{el}(w_x) \frac{d\varphi_{el}}{dx} \qquad (12)$$

Substituting dP^l/dx from Eq. (7) and using $(d\varphi_{el}/dx) = (d\varphi_{el}/dP_x^l)(dP_x^l/dx)$ we obtain

$$\frac{d\varphi_{el}}{dP_x^l} = -\frac{F c_w}{\sigma_{el}(w_x)} \frac{K(w_x)}{\mu} \frac{j_p}{n(w_x) j_p - j_w} \qquad (13)$$

Note that convection of protons with the flow of water has been neglected in Eq. (12). (The effect of a convection term in Eq. (12) was considered in Ref. 16. The net effect due to proton convection is a slight worsening of the effective membrane conductivity.)

The following percolation dependence of proton conductivity will be used for

definiteness,

$$\sigma_{\mathrm{el}}(w_x) = \sigma_{\mathrm{r}} + \sigma_0(w_x - w_{\mathrm{c}})\Theta(w_x - w_{\mathrm{c}}) \tag{14}$$

with percolation threshold w_{c} and residual conductivity σ_{r}. The linear dependence on w_x approximates fairly well with the experimental data for Nafion® 117 [71], although there is of course no ground to expect the exponent to be equal to 1 (in lattice models it is rather close to 2) [88, 91]. The saturation conductivity is $\sigma_{\mathrm{el,s}} = \sigma_{\mathrm{r}} + \sigma_0(w_{\mathrm{s}} - w_{\mathrm{c}})$. Different empirical dependencies can be easily implemented.

Upon increasing the proton current j_{p}, water is dragged electroosmotically through the membrane and leaks out through the cathode gas chamber in a nonstationary fashion. As a result, the total water content decreases to a new stationary value. For each stationary state, a liquid pressure profile, P_x^{l}, is established to compensate the electroosmotic drag. Equation (7) provides the stationary profile w_x, thus, determining the local properties $\sigma_{\mathrm{el}}(w_x)$, $n(w_x)$, and $K(w_x)$. Using these values in Eq. (13), the nonohmic corrections to the current–voltage performance can be calculated.

8.2.2.7.3 **Analytical Solution** Only the essential parts of the solution will be presented in this section. A more detailed consideration including the discussion of analytical expressions for the asymptotic behavior can be found in Refs. 16, 83.

The general solution of Eqs. (7) and (13) is given by

$$j_{\mathrm{p}} = J_{\mathrm{m}} \frac{L}{L - x} g(j_{\mathrm{p}}, r_x) \tag{15}$$

$$U = -\frac{L}{\sigma_0} J_{\mathrm{m}} G(j_{\mathrm{p}}, a),\ U = \varphi_{\mathrm{el}}(L) - \varphi_{\mathrm{el}}(0) \tag{16}$$

with

$$g(j_{\mathrm{p}}, r_x) = \frac{1}{r_1 w_{\mathrm{s}}} \int_{r_x}^{b} \mathrm{d}r \frac{[w(r) - w_{\mathrm{c}}]\Theta(r - r_{\mathrm{c}})\rho(r)}{[n(w(r)) - j_{\mathrm{w}}/j_{\mathrm{p}}]r^2},$$

$$G(j_{\mathrm{p}}, a) = \frac{\sigma_0}{r_1 w_{\mathrm{s}}} \int_{a}^{b} \mathrm{d}r \frac{[w(r) - w_{\mathrm{c}}]\Theta(r - r_{\mathrm{c}})\rho(r)}{\sigma_{\mathrm{el}}(w(r))[n(w(r)) - j_{\mathrm{w}}/j_{\mathrm{p}}]r^2}, \tag{17}$$

valid for $n(w(r))j_{\mathrm{p}} > j_{\mathrm{w}}$ (locally).
r_x is the radius up to which pores are filled in the x-cross section. In pores with these radii the capillary equilibrium prevails. On the cathode side pores are filled up to the radius $r_{\mathrm{L}} = b$, corresponding to the saturation water content. At the anode side we have $r_0 \equiv a$. This solution is valid for $r > r_{\mathrm{c}}$ in all REVs, where r_{c} is the pore radius at the percolation threshold w_{c}. r_1 is the first moment of the pore-radius distribution,

$$r_1 = \frac{1}{w_{\mathrm{s}}} \int_0^{\infty} \mathrm{d}r r \left(\frac{\mathrm{d}w(r)}{\mathrm{d}r} \right) \tag{18}$$

An important membrane parameter J_{m} (in units A cm^{-2}) is defined by

$$J_{\mathrm{m}} = \frac{F\gamma\xi c_{\mathrm{w}} w_{\mathrm{s}}}{4\mu} \frac{r_1}{L} \tag{19}$$

The critical current density j_{pc} is the solution of the implicit equation

$$j_{\mathrm{pc}} = J_{\mathrm{m}} g(j_{\mathrm{pc}}, r_{\mathrm{c}}) \tag{20}$$

implying that the water content at the anode side has dropped to w_{c}.

Fixing j_{p} as a parameter, Eq. (15) gives the profile of r_x, which determines w_x and P_x^{l}. Moreover, calculating a from Eq. (15) by choosing $x = 0$ and using this value in Eq. (16) the current–voltage performance of the membrane is obtained.

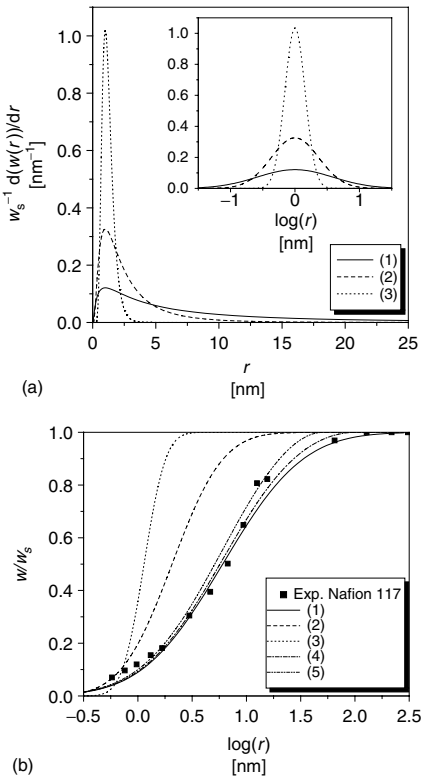


Fig. 5 Pore size distribution function of microporous membranes, parameterization with logarithmic normal distribution, for the parameter sets specified in Table 4. (a) Differential psd, (b) cumulative psd. Experimental data in (b) have been measured by the method of standard porosimetry [44] for Nafion® 117 at 30 °C. Curves correspond to the set of parameters in Table 1.

The difficulty in obtaining this solution directly from Eqs. (15) and (16) is due to the implicit dependencies on the rhs of both equations on j_p. This difficulty disappears, however, if (1) the net water flow is zero, $j_w = 0$, (2) if n is independent of w and, thus, independent of x, or (3) if a δ-function-like psd is assumed. The latter two simplifying assumptions are particularly interesting, since they lead to transparent analytical solutions. However, before discussing these effects in detail, we consider different parameterizations of electroosmotic coefficient n and $w(r)$.

8.2.2.7.4 Parameterization of Membrane Properties

The developed approach will be applied to model membranes whose psds closely resemble those of Nafion® and similar PFSI membranes [60]. A common parameterization of experimental data for ultrafiltration membranes is given by the so-called logarithmic normal distribution [59],

$$\frac{\mathrm{d}w(r)}{\mathrm{d}r} = \frac{w_s}{\Lambda}\left[\exp\left\{-\left(\frac{\log(r/r_m)}{\log s}\right)^2\right\} - \frac{r}{r_{max}}\exp\left\{-\left(\frac{\log(r_{max}/r_m)}{\log s}\right)^2\right\}\right] \quad (21)$$

Tab. 1 Parameters of the psd. Values of the critical current density j_{pc} are calculated on the basis of the membrane parameters specified in Table 2

Set	r_m *[nm]*	s	r_{max} *[nm]*	r_1 *[nm]*	k_s *[nm^2]*	r_c *[nm]*	j_{pc}/J_m	j_{pc} *[A cm^{-2}]*
(1)	1.0	0.15	∞	14.9	1.2×10^3	1.08	0.67	5.0
(2)	1.0	0.30	∞	3.0	1.6×10	0.69	0.65	1.0
(3)	1.0	0.60	∞	1.2	1.5	0.72	0.66	0.4
(4)	1.0	0.15	100	10.8	2.7×10^2	1.05	0.45	2.5
(5)	1.0	0.15	50	8.2	1.2×10^2	1.00	0.40	1.7

In Eq. (21), Λ is a normalization factor, giving the saturation value w_s for the cumulative psd over all pore radii, r_m is a parameter which determines the location of the maximum of the psd, s is a measure of the width of the psd, r_{max} is the maximum pore radius, corresponding to w_s. The psd $dw(r)/dr$ and the cumulative psd $w(r)$, both obtained from Eq. (21), are plotted in Fig. 5 for the parameters in Table 1. The comparison with experimental data for Nafion® 117 reveals, that the parameter set (4) with $r_m = 1$ nm, $s = 0.15$, $r_{max} = 100$ nm gives the best approach to the experimental data [60].

Although this psd exhibits a steep increase for $r < r_m$, it possesses a broad tail in the range of large radii, which could be an indication of large pore defects. Alternatively, it could be interpreted as the manifestation of merging smaller pores into large pores within a soft, deformable polymer matrix. This explanation for the broad psd is consistent with pertinent views on swelling in Nafion®, as described in Ref. 39.

Electroosmotic Coefficient Experimental data on n for Nafion® 117 suggest different options for its parameterization. According to Ref. 4, n is constant over a wide range of w between the water content of a nearly dehydrated membrane (corresponding to $\approx$2 H_2O per SO_3^- group) and the saturation water content in water vapor ($\approx$14–16H_2O per SO_3^- group). Literature values in this range vary between $n = 1.0$ and $n = 1.4$ [4, 86, 87]. In liquid water a different mechanism of water uptake may lead to 50% larger maximum water content at ambient temperature. This leads to larger values $n \approx 2.5$ (at 30 °C, $\approx$22 H_2O per SO_3^- group) [86]. Either the contact with liquid water affects the membrane microstructure in a different way than water vapor, or there is the usual dependence on the membrane prehistory. The controversy of the data of different authors on the w dependence of n is discussed in Ref. 87.

A variant of the model with w-dependent value of n is given in Ref. 16. In the following, we will focus on results obtained for constant n.

8.2.2.7.5 **Solution for Constant *n*** With the assumption of constant n a closed form solution of Eqs. (7) and (13) can be obtained. Defining functions

$$h(r_x) = \frac{1}{r_1 w_s}\int_{r_x}^{b} dr \frac{(w(r)-w_c)\Theta(r-r_c)\rho(r)}{r^2},$$
$$H(a) = \frac{\sigma_0}{r_1 w_s}\int_{a}^{b} dr \frac{(w(r)-w_c)\Theta(r-r_c)\rho(r)}{\sigma_{el}(w(r))r^2} \quad (22)$$

and the parameter

$$j_{pc} = \frac{1}{n}\{j_w + J_m h(r_c)\} \quad (23)$$

the solution for $j_p < j_{pc}$ reads

$$j_p = \frac{1}{n}\left\{j_w + J_m \frac{h(r_x)}{1-(x/L)}\right\} \quad (24)$$

$$U = -\frac{L}{\sigma_0} J_m \frac{j_p}{n j_p - j_w} H(a) \quad (25)$$

Equation (24) determines the water-content profile $w_x \equiv w(r_x)$ for a given current density j_p, which is obtained by inversion

$$x = L\left\{1 - \frac{J_m}{n j_p - j_w} h(r_x)\right\} \quad (26)$$

This relation ensures that there will be no divergence in Eq.(24).

Equation (25) gives the membrane potential drop, $U = \varphi_{el}(L) - \varphi_{el}(0)$, at j_p. Note that in this form, operation parameters (j_p, j_w) have been extracted from the integrals in Eq. (22). Evidently, for known $w(r)$, $\sigma_{el}(w)$, and $K(w)$ the integrations can be carried out explicitly.

For $j_p > j_{pc}$ the current–voltage performance is given by

$$U = \frac{nj_{pc} - j_w}{nj_p - j_w}\frac{j_p}{j_{pc}}U_c - \frac{L}{\sigma_r}\frac{nj_p}{nj_p - j_w}(j_p - j_{pc}) \quad (27)$$

where U_c is the membrane voltage drop at the critical current density j_{pc}, that is, the one at which the water content on the anode side of the membrane drops to the percolation value w_c.

For $j_p > j_{pc}$ the membrane potential drop increases strongly, being inversely proportional to the residual conductivity σ_r, which is usually small compared to σ_0. For $\sigma_r = 0$, j_{pc} is a limiting current density.

8.2.2.8 Exploring the Model

Upon specification of the functions $w(r), n(w), \sigma_{el}(w)$ for a specific PEM sample, the given set of equations provides analytical or numerical solutions for stationary water-content profiles and current–voltage performance. Analytical expressions, which help to rationalize the relevant asymptotic cases (i.e., nonohmic corrections at low current densities, asymptotic behavior near j_{pc}), have been studied in detail [16].

8.2.2.8.1 Effects of the Pore Size Distribution The effect of the psd on the current–voltage plot reveals the relationship between the membrane "capillary portrait" and its performance in a cell. Current–voltage plots in Fig. 6 (for $j_w = 0$) are shown in comparison with the purely Ohmic resistance of the saturated membrane. They suggest that

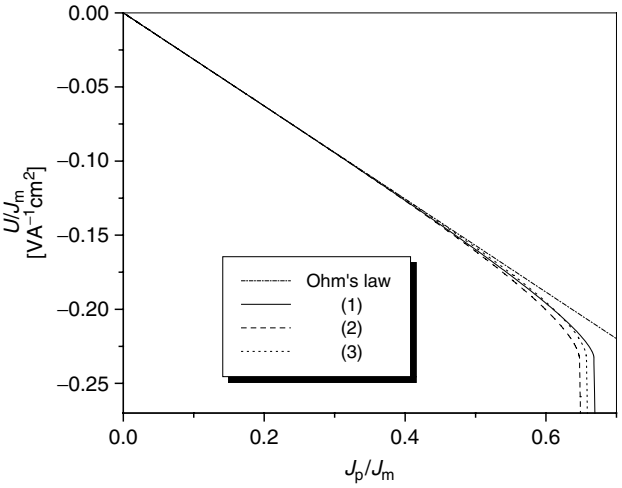


Fig. 6 Effect of the psd on current–voltage performance (for $j_w = 0$). The corresponding psd parameters are listed in Table 1. The other parameters are specified in Table 2. Current densities and voltages have been scaled to $J_m(\propto r_1)$.

Tab. 2 Membrane parameters, used for the calculation of membrane performance under PEFC operating conditions. Parameters are chosen in order to reproduce Nafion® 117 data

Thickness	$L = 0.02$ cm
Conductivity parameter (Eq. 12)	$\sigma_0 = 0.07$ Scm^{-1}
Residual conductivity (Eq. 12)	$\sigma_r = 0.0007$ Scm^{-1}
Saturation water content	$w_s = 0.4$(vol. fraction)
Percolation water content	$w_c = 0.1 w_s$
Surface tension	$\gamma = 0.05$ Jm^{-2}
Viscosity	$\mu \approx 10^{-3}$ kgs^{-1} m^{-1}
Inverse tortuousity	$\xi = 1/3$
Concentration of water	$c_w = 55$ mol l^{-1}
Electroosmotic coefficient	$n = 1.2$

membrane dehydration becomes relevant only in the vicinity of j_{pc}. Below j_{pc} nonlinear corrections to the ideal ohmic performance are small. This is a universal feature of the given model.

Figure 6 accentuates the effect of the width of the psd. The three parameterizations with $s = 0.15, 0.3, 0.6$, labeled as (1), (2), (3) in Table 1 and Fig. 5 are used in Fig. 6. All other parameters of the calculation are specified in Table 2. The width of the psd decreases upon increasing s. The first moments of the psds are, respectively, $r_1 = 14.9, 3.0, 1.2$ nm. In Fig. 6, j_p and U are scaled to $J_m \propto r_1$, defined in Eq. (19). Table 1 indicates that the permeability of the saturated membrane, $K(w_s)$, varies strongly with s. However, the effect of the psd width on membrane performance is accumulated in the scaling parameter J_m, as shown in Fig. 6. Here, the critical current densities are found at $j_{pc}/J_m \approx 0.66, 0.65, 0.67$, practically independent of the psd with this scaling to J_m.

The overall effect of the psd can be narrowed down to the variation of its first moment r_1, which turns up explicitly in the definition of J_m, the cumulative parameter of the model. A larger value of r_1, obtained with larger r_m ($r_1 \propto r_m$) or increased width of the psd (smaller s), gives a larger critical current density j_{pc}, due to the proportionality

$$j_{pc} \propto J_m \propto r_1 \tag{28}$$

Table 1 reveals a significant effect of s on r_1 and, thus, on j_{pc}.

This finding has been further explored in Ref. 83. Due to the demonstrated weak sensitivity of membrane performance to the width of the psd (besides its effect on r_1), it is admissable to use a δ-function like psd in the model,

$$\frac{dw(r)}{dr} = w_s \delta(r - r_1), \quad w(r) = \Theta(r - r_1) \tag{29}$$

which is solely characterized by the first moment r_1 and the saturation water content w_s. This gives an explicit expression for the critical current density,

$$j_{pc} = \frac{1}{n}\left\{ j_w + J_m \left(1 - \frac{w_c}{w_s}\right)\right\} \tag{30}$$

In this equation, essential dependencies of membrane performance on material parameters are explicitly seen. Although derived for an idealized, δ-function psd,

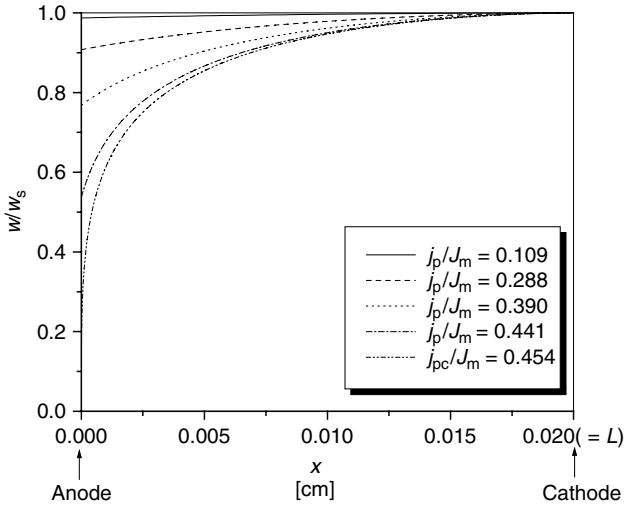


Fig. 7 Stationary water-content profiles at various current densities for the model psd (4), cf. Table 1.

it catches the main qualitative features of membrane behavior: seemingly, the larger the average pore in the membrane, the easier water returns to the anode, and the higher are the current densities that lead to full dehydration of the anode side.

Based on the psd, which gives the best approximation to experimental porosity data for Nafion® 117 (parameterization (4) in Fig. 5), the effects of other parameters on membrane performance have been studied in Refs. 16, 83. We reproduce here (cf. Fig. 7) water-content profiles at various j_p/J_m. The depletion of local water content at the anode side is small for $j_p < j_{pc}$. For $j_p \geq 0.8 j_{pc}$ it becomes remarkable. At $j_p = j_{pc}$ the water content at the anodic membrane boundary drops down to w_c, disturbing the performance. Although the depletion is largely localized in the vicinity of the anode, it is sufficient to limit the current density.

What is the value of J_m, the key parameter of the model? An estimate for Nafion® 117 at $T = 30\,^{\circ}\mathrm{C}$ with the parameters specified in Table 2 and parameterization (4) in Fig. 5 gives $J_m \approx 5\ \mathrm{A\,cm^{-2}}$ (for $j_w = 0$). With this estimate the value $j_{pc} \approx 2.5\ \mathrm{A\,cm^{-2}}$ is obtained. Given this value, membrane dehydration will not be critical for fuel cell operation at $j_p \leq 2.0\ \mathrm{A\,cm^{-2}}$, whereas typical current densities that provide high fuel cell efficiencies are $<1.0\ \mathrm{A\,cm^{-2}}$. Present hydrogen fuel cells commonly use thinner membranes, for example, Nafion® 112 membranes, with a thickness reduced by about a factor 3–4 compared to Nafion® 117. Due to the proportionality $J_m \propto L^{-1}$ critical current densities will be larger by this factor in this membrane. Water management in thinner membranes is generally more facile.

Thus, local dehydration will most likely not be an issue for Nafion® membranes, but that may not be the case for other membranes! For instance, Eq. (28) emphasizes the importance of r_1. If the average pore size is too small, membrane dehydration becomes essential for the fuel cell operation. This is evident from Table 1 by the comparison of j_{pc} values for different psds.

Moreover, a pronounced effect of the maximum pore size is evident from this table. The existence of large pores facilitates water management and thereby gives larger values of j_{pc}.

This conclusion is straightforward as far as the electroosmotic drag coefficient is a weak function of r_1. If it decreases dramatically with the decrease of r_1, the conclusion will depend on the competition between the drag and permeation coefficients. If, however, the whole variation of the drag coefficient is between 1 and 1.4, the conclusions made will stay valid.

8.2.2.8.2 **Modes of Water Management in the Membrane** The model gives hints for water management. Supplying an appropriate flux of water from the anode side to compensate the electroosmotic flux, that is, positive j_w (Eq. 23), will increase the critical current density. Alternatively, one may apply a gradient in gas pressure, $\Delta P^g/L$, as explored in Refs. 16, 83. The effect of this external pressure gradient is superimposed linearly on the internal driving force provided by the capillary pressure gradient.

Indeed, the optimum membrane performance is obtained under saturation conditions, where

$$\frac{d(-P_x^c)}{dx} = 0, \quad \Delta P^g = \Delta P^l, \quad w_x = w_s \tag{31}$$

and

$$U = -\frac{L}{\sigma_s} j_p \text{ for } j_p < j_{ps} = \frac{j_w}{n} + \frac{Fc_w K_s}{n\mu} \frac{\Delta P^g}{L} \tag{32}$$

The parameter j_{ps} is the current density at which membrane dehydration starts. Its value can be controlled via the external parameters j_w and ΔP^g.

This highlights the two basic modes of water management in the membrane. On one hand, maintaining a steady water flow $j_w > nj_p$ (at $\Delta P^g = 0$) is feasible to keep the membrane in a well-hydrated state. This may, however, lead to massive problems with flooding in the cathode compartment. On the other hand, the optimum current–voltage performance can be reached at $j_w = 0$ with

$$\Delta P^g = n \frac{\mu L}{Fc_w K(w_s)} j_p \tag{33}$$

or, if the water produced in the cathode reaction is removed via the anode compartment,

$$\Delta P^g = \left(n + \frac{1}{2}\right) \frac{\mu L}{Fc_w K(w_s)} j_p \tag{34}$$

A numerical estimate for Nafion® 117 with parameters specified in Sect. 8.2.2.8.1 gives

$$\frac{Fc_w K(w_s)}{n\mu L} \approx 1.0 \text{A cm}^{-2}\text{atm}^{-1} \tag{35}$$

Thus, a pressure difference $\Delta P^g \approx 1$ atm would be required to keep the membrane in a completely hydrated state at a current density of 1.0 A cm^{-2}. Again, using a thinner membrane would reduce the required pressure by a factor inversely proportional to the membrane thickness. On the other hand, a strong effect of the psd on this number can be expected according to Table 1: a broader psd and a larger value of r_1 result in considerably larger $K(w_s)$. Correspondingly, a considerably smaller pressure gradient would be required in order to provide conditions of optimum water management, indicated by Eq. (33). If, for instance, r_1 is smaller by a factor 10, then the necessary pressure variation increases by a factor of the order 10^3!

Practical modes of water management include the supply of water at the anode side by humidification of the anodic feed gas, or the application of a pressure across the membrane that creates the backflow from cathode to anode (anode water removal). Different approaches of effective fuel cell water management are discussed, for example, in Ref. 45.

Understanding the laws of membrane dehydration and the possible measures of avoiding it should be taken into account in the regulation of the overall water balance in fuel cell stacks.

8.2.2.9 Generalized Model of Water Backflow

8.2.2.9.1 Generalized Formalism

A generalized membrane transport model, in the form of the ''black box'' models discussed earlier, can be considered in order to compare alternative mechanisms of water backflow in gradients of chemical potential, activity or concentration of water. Each of these gradients can be expressed by a gradient in w. The equation of net water flow is, thus,

$$j_w = n(w_x) j_p - \frac{L}{w_s} J_m^* D(w_x) \frac{dw_x}{dx} \quad (36)$$

Here, J_m^* is a constant with unit of current density and $D(w_x)$ is an effective dimensionless diffusion coefficient. The migration equation reads

$$j_p = -\sigma_{el}(w_x) \frac{d\varphi_{el}(x)}{dx} = -\frac{\sigma_{el}(w_x)}{L J_m^*} \frac{n(w_x) j_p - j_w}{D(w_x)} \frac{d\varphi_{el}}{dw_x} \quad (37)$$

Structural characteristics of the membrane are now incorporated into the transport coefficients $D(w_x)$, $n(w_x)$, $\sigma_{el}(w_x)$. Formally, the generalized model looks like a diffusion model for the interdiffusion of water within the polymer matrix. The effective diffusion model implies water uptake at complete wetting, that is, a homogeneous hydration of the membrane under zero proton current. The attractive feature of this model is its simplicity, since it ignores any explicit elements of phase separation in the membrane, that is, its heterogeneity. It is, thus, free from somewhat artificial assumptions on the three-phase equilibria inside the soft, swelling PEM, typical for the hydraulic model.

For the special case with w-independent n, the solution for $j_p < j_{pc}$ which is given by

$$j_p = \frac{1}{n} \left\{ j_w + J_m^* \frac{1}{1 - (x/L)} \int_{w_x}^{w_s} dw\, D(w) \right\},$$

$$U = -L J_m^* \frac{j_p}{n j_p - j_w} \int_{w_a}^{w_s} dw \frac{D(w)}{\sigma_{el}(w)} \quad (38)$$

(for $j_p > j_{pc}$ it is the same as in Eq. (27))

The critical current density, defined via

$$j_{pc} = \frac{1}{n} \left\{ j_w + J_m^* \int_{w_c}^{w_s} dw\, D(w) \right\} \quad (39)$$

is proportional to the area under the $D(w)$ curve within the limits $w = w_c$ and $w = w_s$. The task of physical models and of ex-situ diagnostics is to determine detailed $D(w)$ dependences.

8.2.2.9.2 Hydraulic Permeation versus Diffusion and Comparison with Experiments

In earlier approaches to the problem of membrane water management, an activity-dependent diffusion coefficient was considered which was transformed into a w-dependent diffusion coefficient, $\mathcal{D}$ (in $cm^2 s^{-1}$) [7]. Within the diffusion approach the nondimensional effective coefficient D is related to the dimensional diffusion constant D (in $cm^2 s^{-1}$) via the identity $J_m^* D = F c_w w_s \mathcal{D}/L$.

In contrast to the diffusion approach, in the previous sections hydraulic permeation was considered as the effective mode of water transport. Transformed to the form of an effective diffusion coefficient the transport coefficient of the latter model becomes

$$D(w) = \frac{1}{r_1}\frac{(w - w_c)\rho(r)}{r^2}\left(\frac{dw(r)}{dr}\right)^{-1} \quad (40)$$

In Fig. 8 the parameterizations for $D(w)$ in the diffusion model (index d), used by Springer et al. [7], and the one obtained with Eq. (40) from the hydraulic permeation model (index h) [16] are compared with each other. They have been scaled in order to give the same value of j_{pc}/J_m^* (same area under the curves). Whereas $D_d(w)$ is practically independent of w, a strong dependence of $D_h(w)$ is observed. In the limit $w \to w_s$, $D_h(w)$ diverges, simply reflecting the fact that pressure driven water transport is possible even if no w-gradient exists (i.e., it is possible to have $\nabla P^l \neq 0$ for $\nabla w = 0$). For $w \to w_c$, $D_h(w)$ smoothly converges to 0. Considering a superposition $J_m^* D = J_d D_d + J_h D_h$ as a likely scenario, Fig. 8 suggests that diffusional transport prevails at low w (strong water–polymer interactions), whereas hydraulic transport dominates in the saturation limit $w \to w_s$, which physically makes sense. This would actually confirm that it is possible to keep the membrane in a homogeneously saturated state by controlling the external pressure gradient ΔP^g, as discussed in Sect. 8.2.2.8.2. Obviously, this mode of water management would not be feasible, if diffusional transport were prevailing in the saturation limit.

Membrane performance characteristics in the "hydraulic" and "diffusion" limits are compared to each other in Fig. 9. Figure 9(a) illustrates that in the diffusion model considerable deviations from the purely ohmic performance of the saturated membrane arise already at small j_p/J_d, well below the critical current density. This is in line with the comparison of the water-content profiles calculated in the diffusion model, Fig. 9(b), with those from the hydraulic permeation model, in Fig. 7. Indeed, membrane dehydration is much stronger in the diffusion model, affecting larger membrane domains at given values of j_p/J_d. Moreover, the profiles exhibit different curvature from those in Fig. 7.

Experimental data by Büchi et al. [66, 92] and Mosdale et al. [93] speak in favor of hydraulic permeation as the basic mode of

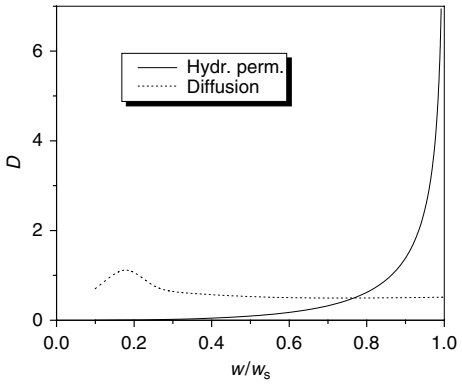


Fig. 8 Parameterizations of effective dimensionless diffusion constants $D(w)$ in the model of hydraulic permeation (cf. Eq. 2.40), using parameterization (4) in Table 1) and in the diffusion model. In the latter, D is obtained from the dimensional diffusion constant $\mathcal{D}$ (in $cm^2 s^{-1}$) via the identity $J_m^* \mathcal{D} = F c_w w_s \mathcal{D}/L$. Absolute values have been adjusted in such a way, that both parameterizations will give the same value of j_{pc}/J_m^*.

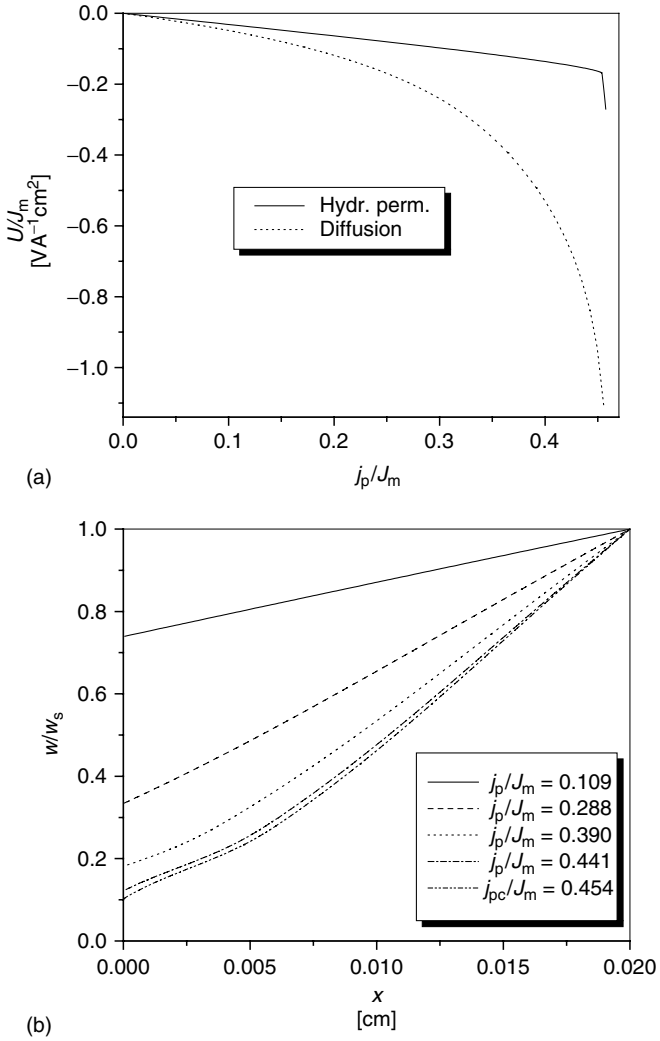


Fig. 9 (a) Current–voltage performances for the model of hydraulic permeation and in the diffusion model. (b) water-content profiles in the diffusion model (compare with Fig. 7).

water backflow, at least under conditions close to saturation. In Refs. 66, 92 it was shown that membrane dehydration is significant only in the thin membrane layer near the anode. The curvature of water-content profiles reported in Ref. 93 resembles the one for the hydraulic permeation model (Fig. 7).

In Fig. 10 the membrane resistance

$$\Re = \int_0^L \mathrm{d}x \frac{1}{\sigma_{\mathrm{el}}(w_x)} \qquad (41)$$

is shown. Experimental data by Büchi and Scherer [94] are compared to values calculated in both model variants. All

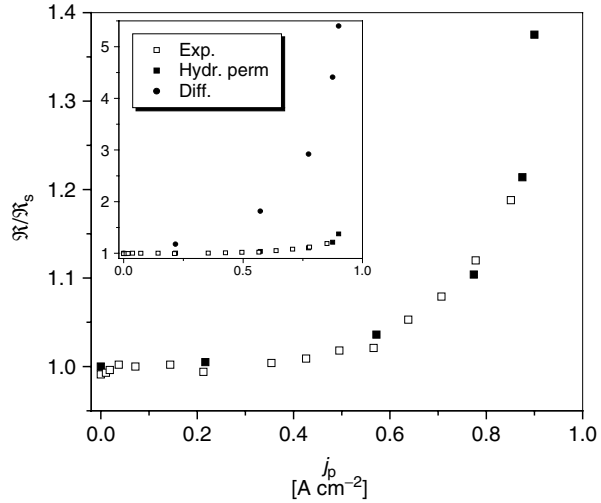


Fig. 10 Membrane resistance in H_2/O_2 fuel cell as a function of proton current density. Experimental data, normalized to the resistance $\Re_s$ of the saturated membrane at various temperatures have been extracted from Ref. 94. They are compared to the values calculated in the hydraulic permeation model (main figure) and to the results of the diffusion model, taken from Ref. 7 (inset).

values are normalized to the resistance of the saturated membrane at various temperatures. A significant increase of resistance is observed in the experimental data only in the vicinity of the critical current density. Experimental data agree well with the results of the hydraulic permeation model, as demonstrated in the main figure. The diffusion model, on the other hand, seems to overestimate the dehydration effect, cf. the inset. This deficiency of the diffusion model is due to the fact that it underestimates the driving force of water backflow in the saturation limit, as indicated by Fig. 9(a).

Recent analysis shows that another parameterization of $D_d(w)$ giving a lower Δ_d at $w \leq 10$, would generate a better agreement of the diffusion model with experimental data [95]. However, any feasible approach has to be based on the proper experimental determination of the corresponding transport coefficient and its incorporation into the appropriate physical model.

A realistic approach should combine in some way elements of hydraulic permeation and diffusion, since generally the ''complex truth'' about water transport in PEM lies, presumably, somewhere between these two limits, with hydraulic permeation mechanism dominating at large and diffusion at small water contents.

Furthermore, which mechanism prevails is also determined by the membrane microstructure and water/polymer interactions. A pronounced hydrophobic/hydrophilic phase separation will result in a well-developed porous structure and, thereby promote hydraulic permeation as the relevant mechanism. In random polymer membranes, which exhibit a smaller extent of ion clustering, water fractions will be more dispersed in the

polymer host. This will fortify the diffusional component of water backflow.

8.2.2.10 Membrane Performance Guidelines

The key concept used in this chapter on membrane performance in PEFC is the perception of present-day PEMs as phase-separated systems. A hydrophobic phase of polymer backbones that provides mechanical stability and a hydrophilic phase of water-containing pathways for proton and water mobility are distinguished.

This notion is supported by a large number of independent experimental data, related to structure and mobility in these membranes. It implies furthermore a distinction of proton mobility in various water environments, strongly bound surface water and liquidlike bulk water, and the existence of water-filled pores as network forming elements. Appropriate theoretical treatment of such systems involves random network models of proton conductivity and concepts from percolation theory, and includes hydraulic permeation as a prevailing mechanism of water transport under operation conditions. On the basis of these concepts a consistent approach to membrane performance can be presented.

The main conclusions for membrane water management are:

Spatial distribution of water: During fuel cell operation, the PEM exhibits an inhomogeneous water distribution. Strong dehydration arises in the interfacial regions close to the anode, whereas the other membrane regions remain in a well-hydrated state, close to saturation. This picture was confirmed experimentally [66, 92].

Critical current density: A critical current density j_{pc} exists at which the total membrane resistance increases strongly. It is reached when the water content w at the anode side drops to the percolation value w_c of the membrane proton conductance. Estimated values of the critical current density $j_{pc} \approx 1\text{–}10\ \mathrm{A\,cm^{-2}}$ for the best currently used membranes lie in the upper range of fuel cell operation, but they could be considerably lower for ''cheaper'' membranes.

The important proportionality $j_{pc} \propto J_m \propto r_1$ was found, revealing the influence of the psd via its first moment r_1. This result is confirmed by experimental data for several membranes, including Nafion® 117 [96]. According to these data, the larger the saturation water content w_s and the smaller the equivalent weight of the membrane, the better the performance characteristics. In Ref. 92 the relation between equivalent weight and dehydration effect was studied by means of in situ resistance measurements on Nafion® 115 and Nafion® 105. Due to the smaller equivalent weight of Nafion® 105, corresponding to larger mean pore radii, no effect of dehydration is observed at $j_p < 1\ \mathrm{A\,cm^{-2}}$, whereas at this current density dehydration is already obvious in Nafion® 115. Of greater importance is the proportionality of j_{pc} to n^{-1}: If n could be suppressed, j_{pc} would be shifted to higher values.

Nonlinear corrections: Nonlinear corrections in current–voltage relations are only relevant in the proximity of j_{pc}. Well below j_{pc} the resistance $\Re_s = L/\sigma_{el,s}$ of the saturated state determines the membrane performance. Above j_{pc} the residual conductivity of dehydrated domains determines the performance.

Membrane thickness: Due to $j_{pc} \propto J_m \propto L^{-1}$ the usage of a thinner membrane will not only reduce $\Re_s = L/\sigma_{el,s}$, but it will also shift the occurrence of the dehydration

effect to larger j_{pc}. Nonohmic corrections are smaller for a thin membrane.

Water management: Water management in the membrane can be controlled via the parameters j_w, the net water flow, and ΔP^g, the gas pressure gradient between cathode and anode. If $j_w > n j_p$ the hydraulic backflow vanishes and w is uniform. The j_w flow may be supplied at the anode and extracted at the cathode. This measure, however, cannot be considered out of relation with the problems with flooding at the cathode. An alternative means of water management control includes the application of an external gradient $\Delta P^g > 0$. The value of j_p at which membrane dehydration becomes relevant is proportional to $\Delta P^g K(w_s)/L$. This indicates that for thinner membranes and membranes with larger permeability a smaller gas pressure gradient is sufficient to provide optimum humidification at a given current density. These suggestions still need further experimental verification.

8.2.3 Structure and Function of Catalyst Layers

8.2.3.1 Catalyst Layer Characterization

The sluggish oxygen reduction reaction (ORR) in the cathode catalyst layer (CCL) induces a major fraction of voltage losses in PEFCs. Due to the requirement of sufficient membrane humidification, discussed in Sect. 8.2.2, and the limitation it imposes on the working temperature range ($<100\,^{\circ}C$) only platinum-based catalysts can provide acceptable reaction rates. Platinum, however, is costly, and its resources are limited.

In view of these restrictions the amount of Pt in catalyst layers of fuel cells should be minimal and, moreover, this amount should be optimally utilized. Systematic understanding of structure vs. function relationships is, thus, crucial for the development of improved, less expensive electrodes.

The essential ingredients of the catalyst layer are an electronically conducting matrix of carbon grains, Pt catalyst particles supported on carbon and a proton-conducting network of well-humidified PFSI. In addition, Teflon (PTFE) may be added as a binder and hydrophobizing agent.

A number of different methods exist for the production of catalyst layers [97–102]. They use variations in composition (contents of carbon, Pt, PFSI, PTFE), particle sizes and pds of highly porous carbon, material properties (e.g., the equivalent weight of the PFSI) as well as production techniques (sintering, hot pressing, application of the catalyst layer to the membrane or to the gas-diffusion layer, GDL) in order to improve the performance. The major goal of electrode development is the reduction of Pt and PFSI contents, which account for substantial contributions to the overall costs of a PEFC system. Remarkable progress in this direction has been achieved during the last decade [99, 100]. At least on a laboratory scale, the reduction of the Pt content from 4.0 to 0.1 mg cm^{-2} has been successfully demonstrated.

Currently utilized catalyst layers are three-phase composites. A typical structure is sketched in Fig. 11, depicting the interpenetrating functional phases. A solid phase of carbon grains serves as the supplying network for the electrons. Similarly, the electrolyte network, embedded in the carbon structure, provides the pathways for the protons. Open pores in the composite comprise a network for the diffusion of gases. Carbon grains are loaded with Pt particles with sizes in the range 1–10 nm. The key to an excellent electrode performance with low Pt loadings is

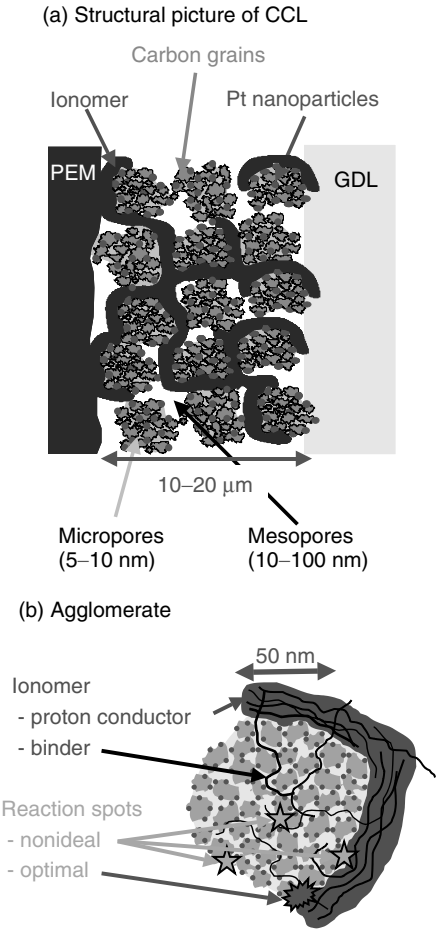


Fig. 11 Schematic picture of the cathode catalyst layer and its composition, exhibiting the different functional parts. The typical catalyst layer thickness is $\ell \approx 10-20\,\mu m$.

the creation of a large electroactive contact area between carbon/Pt particles and electrolyte particles.

Because of the intricate composition, various transport modes contribute to the supply of O_2 to the reaction sites, involving viscous flow, Knudsen diffusion and ordinary molecular diffusion [103]. The relative importance of the distinct mechanisms depends decisively on the psd. Dominating transport though micropores favors Knudsen flow, whereas molecular diffusion with considerably larger diffusion constants will prevail in macropores. A critical pore radius for the transition between Knudsen and molecular diffusion is ~30 nm [104].

Moreover, due to a net water flow toward the cathode and the production of water in it, oxygen diffusivity will be a function of the current density. At larger current densities larger amounts of water will accumulate within the cathode, thereby, hampering gaseous transport. Modeling approaches that incorporate the important issue of liquid water formation and partial saturation in CCLs have been developed only recently. They reveal a key role of the CCL in regulating the fuel cell water fluxes.

According to the porosity data of Uchida et al. [102] the matrix of carbon grains (20–40 nm) forms an agglomerated structure with a bimodal psd. Primary pores (micropores, 5–40 nm) exist within agglomerates, between the carbon grains. Larger, secondary pores (macropores, 40–200 nm) form the pore spaces between agglomerates. The relation between the relative pore volume fractions of the two pore types depends on the contents of PFSI and PTFE. Due to their molecular size these components are not able to penetrate micropores. They affect only the macropore volume. The experimental study revealed that an increased PFSI content leads to a decrease of the macropore volume fraction. The opposite effect was found for PTFE.

Two key aspects of CCL operation will be subsequently considered:

1. understanding the interplay of physical processes that determine the operation,

2. rationalizing the effect of the composition on performance.

The theory of "gas-diffusion electrodes" has a long history [11, 12, 105, 106]. Specifically, catalyst utilization and specific effective surface area in composite electrodes were always in the focus of attention (see, e.g., Ref. 13 and the articles quoted therein). A comprehensive review of the theory and models of ionic-into-electronic current transformation in two- and three-phase distributed electrodes can be found in Ref. 14.

An appropriate model of CCL operation should account for the interdiffusion of oxygen, nitrogen, and water vapor, migration of protons and electrons, and kinetics of the electrochemical reaction. The pertinent theory was developed in Refs. 8, 9, 15. A model similar to that was more recently studied analytically [17, 107, 108], and we will dwell on these transparent closed form results. We will see that having such solutions at hand helps in revealing the reserves for optimization of the structure and function of the catalyst layers.

8.2.3.2 Model of Catalyst Layer Performance

For modeling purposes, the CCL is considered an effective medium of thickness l with homogeneous distributions of the various components. Moreover, homogeneous boundary conditions in the interfacial layers between catalyst layer and GDL on one side and membrane on the other side are assumed. Therefore, here the transport and reaction processes in the layer depend solely on the coordinate x with the origin at the PEM| CCL interface (cf. Fig. 11).

Based on the assumption of good electronic conductivity of the carbon/catalyst phase, ohmic losses due to electronic transport can be neglected, which is a good approximation unless the carbon network is too close to the percolation threshold. The equation of oxygen diffusion can be written as

$$\frac{\mathrm{d}p(\chi)}{\mathrm{d}\chi} = \frac{j_0 - j_\mathrm{p}(\chi)}{I}, \quad I = \frac{4F\bar{P}_{\mathrm{O}_2}\mathcal{D}}{RTl}, \quad \chi = \frac{x}{l} \tag{42}$$

Here, the oxygen partial pressure p is normalized to the absolute O_2-partial pressure $\bar{P}_{\mathrm{O}_2}$ at the interface between catalyst layer and GDL (at $\chi = 1$), $p = P_{\mathrm{O}_2}/\bar{P}_{\mathrm{O}_2}$. D is an effective oxygen diffusion constant (in $\mathrm{cm^2 s^{-1}}$). $j_\mathrm{p}(\chi)$ is the local proton current density (in $\mathrm{A\,cm^{-2}}$) and $j_0 = j_\mathrm{p}(\chi = 0)$ is its value at the interface with the membrane, where j_0 is equal to the total current density through the cell.

The parameter I (in $\mathrm{A\,cm^{-2}}$) is a characteristic current density of oxygen diffusion. The factor 4 in the definition of I accounts for the number of protons that are consumed by one oxygen molecule in the cathode reaction. Equation (42) is based on Knudsen diffusion as the prevailing mechanism of oxygen transport in a predominantly microporous catalyst layer, that is, oxygen molecules are assumed to collide more frequently with pore walls than with other molecules [104]. A variant which, instead, considered molecular diffusion as the prevailing mode of oxygen transport was studied in Ref. 17. This would be more realistic in a layer with a considerable porosity due to macropores. Both variants give similar results. Analytical expressions are, however, simpler in the variant presented here.

Proton transport in the layer is described by the migration equation

$$\frac{\mathrm{d}\eta(\chi)}{\mathrm{d}\chi} = -\frac{1}{\sigma} j_\mathrm{p}(\chi), \quad \sigma = \frac{\sigma_\mathrm{el}}{l} \tag{43}$$

where $\eta(\chi)$ is the local electrode potential, defined here simply as the difference between the potentials of electrolyte, φ_{el}, and carbon/catalyst phase, φ_c, $\eta(\chi) = \varphi_{el}(\chi) - \varphi_c(\chi)$ (skipping over details in definitions of potentials in the interfacial region). Under the assumption of high electronic conductivity of the carbon/catalyst phase, this phase can be considered as equipotential, deliberately fixing $\varphi_c(\chi) = 0$. Variations in the electrode potential η are, thus, equal to variations in φ_{el} alone. σ_{el} is the proton conductivity of the layer.

The net rate of electrochemical current generation in an electrode section is given by the Butler–Volmer equation, a fundamental relation in electrode kinetics [109],

$$\frac{\mathrm{d}j_p(\chi)}{\mathrm{d}\chi} = -i^* \left[p(\chi) \exp\left(\frac{\alpha_c n_t F \eta(\chi)}{RT}\right) - \exp\left(-\frac{(1-\alpha_c) n_t F \eta(\chi)}{RT}\right)\right] \qquad (44)$$

Here, n_t is the number of electrons that are transferred in the rate-determining reaction step. At a Pt| PFSI interface $n_t = 1$ has been identified [99, 110]. The standard exchange current density (in A cm^{-2})

$$i^* = 2F S_0 l k^* (c_{H^+})^\nu \bar{c}_{O_2}$$

is an effective reaction rate parameter, determined by the externally provided reactant concentrations. The factor 2 accounts for the number of electrons that are transferred in one elementary reaction. Moreover, i^* is proportional to the total effective surface S_0 of the electroactive Pt| PFSI contact area per unit volume (unit cm^{-1}) and to the reaction rate constant k^*. The dimensionless property $\xi = S_0 l$ corresponds, thus, to the real-to-apparent surface area ratio.

Due to the assumption of electroneutrality in any REV of the layer, the proton concentration c_{H^+} and, thus, the term $(c_{H^+})^\nu$ is assumed to be constant over the whole sample. Therefore the factor in front of the second exponential in Eq. (44) is one. $\bar{c}_{O_2}$ is the external oxygen molar concentration at $\chi = 1$, corresponding to $\bar{P}_{O_2}$ via the ideal gas law. Following Refs. 99, 110, 111 the cathodic oxygen reduction is assumed to be a first-order reaction, the current density depending linearly on the local O_2 partial pressure.

The apparent transfer coefficient of the cathodic reaction, α_c, is a measure of the sensitivity of the transition state to the drop in electrostatic potential between electrolyte and metal [109, 112]. According to Ref. 113, it is $\alpha_c = 0.75$ for the O_2 reduction on Pt in aqueous acid electrolytes. In Ref. 111 the value $\alpha_c = 1.0$ was reported instead. Since the cathodic reaction is a complex multistep process, it might follow several reaction pathways, and the competition between them is affected by the operation conditions (η, p, T). Therefore, different values of α_c have been reported in different regimes of operation. Although in the simple reactions the transfer coefficient is a microscopic characteristic of the elementary act [112], for complex multistage reactions in fuel cell electrodes it is rather an empirical parameter of the model. The dependence of effective α for methanol oxidation on the catalyst layer preparation was recently studied [114].

At small electrode potentials, $\eta \leq b/3$, only small amounts of reactants will be consumed and it is, thus, reasonable to put $p(\chi) = 1$. Series expansions of the exponentials in Eq. (44) lead to a linear relation

$$\frac{\mathrm{d}j_p(\chi)}{\mathrm{d}\chi} = -i^* \frac{\eta(\chi)}{\tilde{b}}, \tilde{b} = \frac{RT}{n_t F} \qquad (45)$$

whereas at $\eta \geq 2b$ it is conventionally approximated by the Tafel expression,

$$\frac{\mathrm{d}j_\mathrm{p}(\chi)}{\mathrm{d}\chi} = -i^* p(\chi) \exp\left(\frac{\eta(\chi)}{b}\right), \quad b = \frac{RT}{\alpha_\mathrm{c} n_\mathrm{t} F} \tag{46}$$

where b is the Tafel parameter.

The basic equations of catalyst layer operation, Eqs. (42–46), are valid under the assumption of isothermal, stationary conditions. Furthermore, variations of the water vapor partial pressure are neglected. The water content in the PFSI fractions and the corresponding proton conductivity are, therefore, independent of χ. Upon proceeding along χ, starting at $\chi = 0$ with $j_\mathrm{p}(\chi = 0) = j_0$, proton current is gradually converted into O_2 flux $j_{O_2} = (j_\mathrm{p}(\chi) - j_0)/4$. At $\chi = 1$ the transformation is complete, $j_\mathrm{p} = 0$, since no protons are admitted to pass the interface to the GDL.

The voltage loss U_c incurred in the catalyst layer, is defined as the difference between local electrolyte potential at $\chi = 0$ and local metal phase potential at $\chi = 1$, $U_\mathrm{c} = \varphi_\mathrm{el}(\chi = 0) - \varphi_\mathrm{c}(\chi = 1)$. However, since $\varphi_\mathrm{c}(\chi) \equiv 0$ throughout the layer, U_c is given by the local electrode potential at $\chi = 0$,

$$U_\mathrm{c} = \eta(\chi = 0) = \eta_0 \tag{47}$$

The boundary conditions imposed on the three independent variables p, j_p, and η at $\chi = 1$ are $p = 1, \eta = \eta_1, j_\mathrm{p} = 0$. Fixing the free parameter η_1, the corresponding value of j_0 can be obtained from the self-consistency equation $j_\mathrm{p}(\chi = 0) = j_0$. The complete solution then gives the profiles $p(\chi)$, $j_\mathrm{p}(\chi)$, and $\eta(\chi)$, as well as the polarization relation $\eta_0(j_0)$.

The resulting curves depend on externally controlled operation conditions $(\bar{c}_{O_2}, T)$ and on three composition-dependent parameters:

1. I, the characteristic current density of oxygen diffusion
2. σ, the proton conductance per geometric surface area
3. i^*, the exchange current per geometric surface area.

Generally, the nonlinear system of equations may only be solved numerically. However, analytical solutions exist in certain limiting cases, which will be considered first.

8.2.3.3 Close to Open-circuit Potentials

Cathode potentials in the range $\eta_0 < b/3$ are of limited practical importance for PEFC operation. Nevertheless, it is insightful to consider this case first, since it illuminates a number of basic aspects of electrode performance. In this regime only small amounts of oxygen are consumed, i.e., $p \equiv 1$. A simplified, textbook set of equations [105, 106] emerges

$$\frac{\mathrm{d}j_\mathrm{p}(\chi)}{\mathrm{d}\chi} = -i^* \frac{\eta(\chi)}{\tilde{b}}, \quad \frac{\mathrm{d}\eta(\chi)}{\mathrm{d}\chi} = -\frac{1}{\sigma} j_\mathrm{p}(\chi), \quad \tilde{b} = \frac{RT}{n_\mathrm{t} F} \tag{48}$$

with the well-known solution

$$j_\mathrm{p}(\chi) = \frac{\sinh\left(\frac{l}{l_\mathrm{p}}(1-\chi)\right)}{\sinh\left(\frac{l}{l_\mathrm{p}}\right)} j_0, \quad \eta(\chi) = \frac{1}{\sigma_\mathrm{p}} \frac{\cosh\left(\frac{l}{l_\mathrm{p}}(1-\chi)\right)}{\sinh\left(\frac{l}{l_\mathrm{p}}\right)} j_0 \tag{49}$$

The polarization relation is, therefore, given by

$$\eta_0 = \frac{j_0}{\sigma_p} \coth\left(\frac{l}{l_p}\right) \tag{50}$$

The corresponding linear electrode polarization resistance is

$$\Re_c = \frac{\eta_0}{j_0} = \frac{1}{\sigma_p} \coth\left(\frac{l}{l_p}\right) \tag{51}$$

Here, the parameters

$$\sigma_p = \left(\frac{\sigma i^*}{\tilde{b}}\right)^{1/2} \tag{52}$$

and

$$\frac{l_p}{l} = \left(\frac{\sigma \tilde{b}}{i^*}\right)^{1/2} \tag{53}$$

are introduced, where l_p is the so-called reaction-penetration depth. Equation (49) indicates that the profile of $j_p(\chi)/j_0$, which suggests a pattern of the catalyst utilization, is independent of j_0. It is solely determined by l_p. A reaction-penetration depth $l_p > l$ gives homogeneous catalyst utilization, whereas for $l_p \ll \ell$ only a small fraction of the catalyst close to the membrane is used for the reaction, since a small proton conductivity inhibits protons from penetrating the layer. The value of l_p may serve as an orientation for an optimal catalyst layer thickness, which would provide a compromise between high catalyst utilization and small η_0. In order to reduce η_0, Eq. (50) suggests that σ_p, an effective conductivity of the catalyst layer, should be large. This can be achieved with large electrolyte conductivity σ_{el}, large specific catalyst surface S_0, large reaction rate k^*, and large oxygen concentration $\bar{c}_{O_2}$.

8.2.3.4 High Electrode Potentials

8.2.3.4.1 Dimensionless Form of Basic Equations At $\eta_0 \geq 3b$ the rate of current generation is well approximated using the Tafel law given by Eq. (46). It is convenient to rewrite the basic set of equations in an entirely dimensionless form [17],

$$\frac{dp}{d\zeta} = -(1 - i) \tag{54}$$

$$\frac{d\Gamma}{d\zeta} = g\Gamma i \tag{55}$$

$$\frac{di}{d\zeta} = p\Gamma \tag{56}$$

where the three independent variables i, p, and Γ are functions of ζ, with the notations

$$\zeta = \frac{j_0}{I}(1 - \chi),\ i = \frac{j_p}{j_0},\ \Gamma = \frac{i^* I}{j_0^2} \exp\left(\frac{\eta}{b}\right) \tag{57}$$

Now it is obvious that for fixed l the solution is determined by a single parameter,

$$g = \frac{I}{\sigma b} \tag{58}$$

which allows us to distinguish between diffusion limited ($g \ll 1$), conductivity limited ($g \gg 1$) and mixed ($g = 1$) cases. From the definition of Γ we immediately retain

$$\eta = b \ln \Gamma + 2b \ln \frac{j_0}{I} - b \ln \frac{i^*}{I} \tag{59}$$

Equations (54–56) and Eq. (59) reveal three important features of the solution. First, since the system of Eqs. (54–56) is invariant with respect to i^*, the explicit contribution of i^* to η is incorporated in the logarithmic term in Eq. (59). Second, the appearance of a term $2b \ln(j_0/I)$ with twice the Tafel slope of the pure reaction is a universal result, not due to any specific parameter values, following only

from the structure of the equations. Under certain conditions this term can give the dominating contribution to η. Third, since g is invariant with respect to l, the effect of l on the catalyst layer performance can be easily rationalized from Eqs. (54–56). A change of thickness from l to $\tilde{l} = \kappa l$, at a simultaneous change from j_0 to $\tilde{j_0} = j_0/\kappa$ will lead to a shift of $\tilde{\eta_0} = \eta_0 - 2b \ln \kappa$. These features help to determine the thickness dependence of voltage losses incurred by the cathode, as discussed in Sect. 8.2.3.5.3.

8.2.3.4.2 **General Solution** For fixed boundary conditions

$$p(\zeta = 0) = 1,\ i(\zeta = 0) = 0,\quad \Gamma(\zeta = 0) = \Gamma_1 \tag{60}$$

the profiles in $p(\zeta)$, $\Gamma(\zeta)$, and $i(\zeta)$ can be calculated. Matching the self-consistency condition

$$i\left(\zeta = \frac{j_0}{I}\right) = 1 \tag{61}$$

gives j_0 and,

$$\Gamma_0 = \Gamma\left(\zeta = \frac{j_0}{I}\right) \tag{62}$$

Finally, the cathode potential is calculated from

$$\eta_0 = b \ln \Gamma_0 + 2b \ln \frac{j_0}{I} - b \ln \frac{i^*}{I} \tag{63}$$

8.2.3.4.3 **Limit of Fast Oxygen Diffusion** The case of fast oxygen transport through the catalyst layer is of practical interest, since nowadays thin catalyst layers (5–10 μm) are usually used. One may expect the oxygen concentration to vary only moderately across such a thin layer. The catalyst layer performance is then governed by Eqs. (43) and (46), with fixed normalized oxygen pressure $p \equiv 1$. These two equations are equivalent to a Poisson–Boltzmann-like expression,

$$\frac{\mathrm{d}^2\eta(\chi)}{\mathrm{d}x^2} = \tilde{b}\left(\frac{l}{l_\mathrm{p}}\right)^2 \exp\left(\frac{\eta(\chi)}{b}\right) \tag{64}$$

which can be exactly integrated to give a solution in parametric form

$$\eta(\chi) = \eta_1 - 2b \ln\left\{\cos\left[\sqrt{\frac{\alpha_\mathrm{c}}{2}}\frac{l}{l_\mathrm{p}}(1-\chi)\exp\left(\frac{\eta_1}{2b}\right)\right]\right\} \tag{65}$$

$$j_\mathrm{p}(\chi) = \sqrt{2\sigma b i^*}\exp\left(\frac{\eta_1}{2b}\right) \times \tan\left[\sqrt{\frac{\alpha_\mathrm{c}}{2}}\frac{l}{l_\mathrm{p}}(1-\chi)\exp\left(\frac{\eta_1}{2b}\right)\right] \tag{66}$$

where the working point is fixed by specifying a value η_1, the electrode potential at $\chi = 0$. The potential losses in the catalyst layer are obtained by putting $\chi = 0$ in these equations and calculating the corresponding $\eta_0 = \eta(0)$ and $j_0 = j_\mathrm{p}(0)$ at the given η_1. Note that the factor $\sqrt{2\sigma b i^*}$ is independent of l, since $\sigma \propto l^{-1}$ and $i^* \propto l$.

An alternative way toward this solution was pursued in Ref. 108. Differentiating Eq. (46) with respect to χ and substituting $\mathrm{d}\eta/\mathrm{d}\chi$ from Eq. (43) an expression leads to

$$\frac{\mathrm{d}}{\mathrm{d}\chi}\left(\frac{\mathrm{d}j_\mathrm{p}}{\mathrm{d}\chi} + \frac{1}{2\sigma b}{j_\mathrm{p}}^2\right) = 0 \tag{67}$$

The latter equation indicates the spatial invariance of the term in brackets,

$$\frac{\mathrm{d}j_\mathrm{p}}{\mathrm{d}\chi} + \frac{1}{2\sigma b}{j_\mathrm{p}}^2 = \left.\frac{\mathrm{d}j_\mathrm{p}}{\mathrm{d}\chi}\right|_{\chi=0} + \frac{1}{2\sigma b}{j_0}^2 \tag{68}$$

Further integration of this equation subject to the boundary condition $j_\mathrm{p}(\chi = 1) = 0$ yields a single equation, which relates j_0

and η_0 [108],

$$\frac{j_0}{2\sigma b} = \sqrt{\frac{\alpha_c}{2}\left(\frac{l}{l_p}\right)^2 \exp\left(\frac{\eta_0}{b}\right) - \left(\frac{j_0}{2\sigma b}\right)^2} \times \tan\left(\sqrt{\frac{\alpha_c}{2}\left(\frac{l}{l_p}\right)^2 \exp\left(\frac{\eta_0}{b}\right) - \left(\frac{j_0}{2\sigma b}\right)^2}\right) \tag{69}$$

It is not possible to resolve this transcendental equation for the explicit $\eta_0(j_0)$ dependence. However, two limiting cases can be conveniently studied: in the limit $j_0 \ll 2\sigma b$ the simple Tafel kinetics is recovered,

$$\eta_0 = b \ln\left(\frac{j_0}{i^*}\right), \quad j_0 \ll 2\sigma b \tag{70}$$

whereas under conditions of severe proton transport limitations, $j_0 \gg 2\sigma b$, the electrode potential reveals a characteristic double Tafel-slope dependence

$$\eta_0 = 2b \ln\left(\frac{j_0}{\sqrt{2\sigma b i^*}}\right), \quad j_0 \gg 2\sigma b \tag{71}$$

Note that according to Eq. (65) the latter limit is attained when the local electrode potential η_1 at $\chi = 1$ approaches a j_0 - independent maximum value

$$\eta_{1,\max} = 2b \ln\left(\frac{\pi}{\sqrt{2\alpha_c}}\frac{l_p}{l}\right) \tag{72}$$

whereas η_0 increases according to Eq. (71). The corresponding potential drop in the electrolyte phase is

$$\eta_0 - \eta_1 = 2b \ln\left(\frac{j_0}{\pi\sigma b}\right), \quad j_0 \gg 2\sigma b \tag{73}$$

Figure 12 shows the current–voltage plot of the catalyst layer in log–log coordinates. The transition from $j_0 = i^* \exp(\eta_0/b)$ at small current densities to $j_0 = \sqrt{2\sigma b i^*}\exp(\eta_0/2b)$ at large current densities is clearly seen in this plot. Near $j_0 = 2\sigma b$ the transition region described by Eq. (69) couples the two limiting straight lines.

8.2.3.4.4 Limit of High Proton Conductivity

In the case of negligible proton transport limitations, that is, $\eta(\chi) = \eta_0$, but possible oxygen diffusion limitations the catalyst layer performance is determined by

$$\frac{d^2 p(\chi)}{d\chi^2} = \left[\frac{i^*}{I}\exp\left(\frac{\eta_0}{b}\right)\right] p(\chi) \tag{74}$$

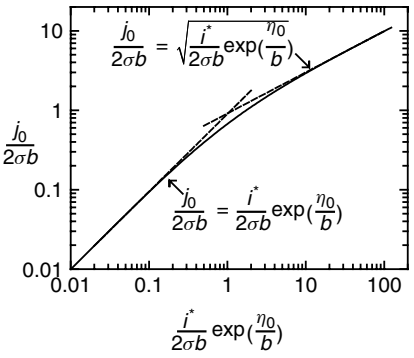


Fig. 12 Log–log plot of normalized current density versus exponent of normalized voltage losses incurred by the cathode catalyst layer ($U_c = \eta_0$), in the limit of fast oxygen diffusion (Sect. 8.2.3.4.3). This representation reveals the transition from the simple Tafel kinetics at $j_0 \ll 2\sigma b$ to the double Tafel-slope dependence at $j_0 \gg 2\sigma b$.

Its solution gives the profiles

$$p(\chi) = \sinh(\Upsilon_0(1-\chi))[\coth(\Upsilon_0(1-\chi)) - \tanh(\Upsilon_0)],\ \Upsilon_0 = \sqrt{\frac{i^*}{I}}\exp\left(\frac{\eta_0}{2b}\right),$$
$$j_p(\chi) = I\Upsilon_0[\tanh(\Upsilon_0) - \tanh(\Upsilon_0) \times \cosh(\Upsilon_0(1-\chi)) + \sinh(\Upsilon_0(1-\chi))] \quad (75)$$

and

$$j_0 = I\Upsilon_0\tanh(\Upsilon_0) \quad (76)$$

However, for $\eta_0 < b\ln(I/i^*)$ ($\Upsilon_0 < 1$), oxygen diffusion limitations can be neglected as well, and one may put $p \equiv 1$ throughout the layer. As a result, the reaction determined Tafel law, Eq. (46), will control the performance. But for $\eta_0 > b\ln(I/i^*)$ ($\Upsilon_0 > 1$) oxygen diffusion is critical and the performance is given by

$$\eta_0 = 2b\ln\left(\frac{j_0}{\sqrt{i^*I}}\right) \quad (77)$$

with

$$p(\chi) = \exp\left(-\frac{j_0}{I}(1-\chi)\right),$$
$$j_p(\chi) = j_0\left[1 - \exp\left(-\frac{j_0}{I}(1-\chi)\right)\right] \quad (78)$$

These profiles reveal the thickness of the effective layer, adjacent to the backing layer,

$$\delta_{\text{eff}} = \frac{I}{j_0}l \quad (79)$$

in which the major fraction of the current density is generated.

8.2.3.4.5 **The Double Tafel Slope as a Signature of Transport Limitations** The kind of transport limitations that prevail depend on the structure of the catalyst layer. If it has insufficient porosity, but a continuous network of polymer electrolyte, it will have severe gas transport limitations, but good proton transport and vice versa. The dominance of one transport limitation, be it proton transport (Sect. 8.2.3.4.3) or oxygen diffusion (Sect. 8.2.3.4.4), results in a logarithmic current–voltage plot with double Tafel slope at large j_0. The characteristic current densities are $\sqrt{2\sigma bi^*}$ or $\sqrt{i^*I}$, respectively, which are both independent of l, demonstrating that in the regime with double Tafel-slope signature the electrode performance is l invariant.

8.2.3.5 General Performance Characteristics

There are new features of performance if both transport limitations are present at the same time. Understanding them takes more effort and a more general consideration, which, however, also leads to clear-cut conclusions on the current–voltage performance and the profile of $p(\chi)$ and $j_p(\chi)/j_0$.

8.2.3.5.1 **Basic Regimes of Performance** These could be classified subject to the values of the parameter g [17]: (1) $g \ll 1$ (strong oxygen diffusion limitations, fast proton transport), (2) $g \approx 1$ (mixed kinetics), (3) $g \gg 1$ (fast oxygen diffusion, strong proton transport limitations).

Results for case (2) are shown in Fig. 13. In polarization curves three regimes of current density can be distinguished, which are indicated in Fig. 13(a): (1) the kinetic regime, $j_0 \ll I$, with a simple Tafel dependence, $\eta_{\text{kin}} = b\ln(j_0/i^*)$, determined by the kinetic parameter i^*, (2) the intermediate regime, $j_0 \approx I$, with a dominating double Tafel-slope behavior $\eta_{\text{int}} = 2b\ln(j_0/I)$, (3) the oxygen exhaustion regime, $j_0 \gg I$, in which all oxygen is consumed in an active sublayer of

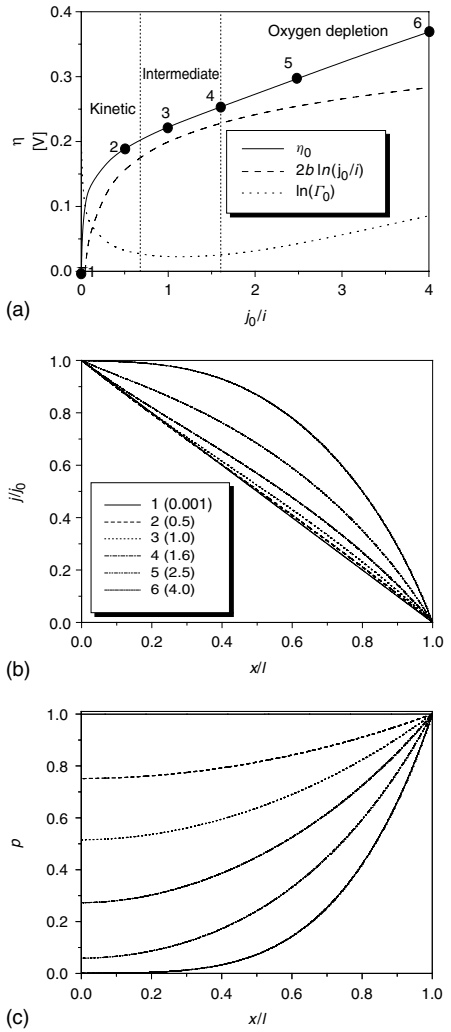


Fig. 13 Dimensionless plot of polarization characteristics of the catalyst layer and profiles of current density and pressure along the thickness coordinate χ. Here, these properties are depicted in the mixed regime, when proton and oxygen transport limitations are of equal importance, $I = \sigma b$ ($g = 1$). (a) The potential drop η_o and particular contributions according to Eq. (81) ($2b\ln(j_0/I)$, $b\ln\Gamma_c$) are shown, helping to distinguish the three indicated regimes of performance. (b) The profiles of j_p/j_0 and (c) p are depicted at the points 1–6 marked in (a). Local distributions in electrode potential, which are also calculated in the model, are not shown here. The general trend observed from them is an increase towards the membrane interface. At point 6 the total increase is $\approx 2b$.

thickness δ_{eff}, defined in Eq. (79), adjacent to the backing layer. Voltage losses due to proton transport show a characteristic slope proportional to $1/\sigma$ in the oxygen depletion regime, incurred in the remaining inactive sublayer. It is to be noted that the distinction of the three regimes of performance is an idealization, since in reality their boundaries can be defined only with a certain tolerance. However, their approximate identification in experimental electrode polarization curves will provide useful estimates of the fitting parameters i^*, I, σb.

8.2.3.5.2 **Effect of Oxygen Depletion** At various values of j_0, labeled with dots in Fig. 13(a), local distributions of $j_p(\chi)/j_0$ and $p(\chi)$ are shown in Fig. 13(b) and (c). In the kinetic regime (points 1 and 2) oxygen homogeneously penetrates the whole catalyst layer and reaction rates are

likewise distributed homogeneously. The intermediate regime (points 3 and 4) is characterized by a considerable decrease of p near the membrane. Corresponding distributions of $j_p(\chi)/j_0$ are slightly nonlinear, indicating a modest increase of reaction rates near the backing layer. Nevertheless, in the intermediate regime all the catalyst particles are used for reactions, i.e. there is no definite loss in catalyst utilization. Finally, in the oxygen depletion regime (points 5 and 6) both profiles are strongly nonlinear. In this regime, due to the poor oxygen diffusivity, practically all the supplied O_2 is consumed in an active sublayer $(1 - \delta_{\text{eff}}/l) < \chi < 1$, where δ_{eff} is inversely proportional to j_0 according to Eq. (79). The remaining sublayer on the membrane side is not used for reactions. The electrode potential in the inactive region $0 < \chi < (1 - \delta_{\text{eff}}/l)$ is given by

$$\eta(\chi) \approx \eta_{\chi=1-\delta_{\text{eff}}/l} + \frac{j_0}{\sigma}(1 - \chi - \delta_{\text{eff}}/l) \tag{80}$$

These voltage losses are most pronounced for low proton conductivity, that is, small σ. The ohmic contribution ($\propto \sigma^{-1}$) results from a simple fact: if there is a layer close to the membrane with no oxygen, the protons must pass this layer (without any reaction) to reach another layer, close to the backing layer, where there is oxygen to react with. Actually, at current densities $j_0 \geq I$ the catalyst layer performs as if it had one flexible boundary at $\chi = 1 - \delta_{\text{eff}}/l$. While the boundary on one side, $\chi = 1$, remains fixed, the boundary at $\chi = 1 - \delta_{\text{eff}}/l$ moves away from the membrane interface with increasing current density. The electrode potential for $\delta_{\text{eff}} \ll l$ is

$$\eta_0 \approx b \ln \Gamma_c + 2b \ln \frac{j_0}{I} + \frac{j_0}{\sigma} - b \ln \frac{i^*}{I} \tag{81}$$

with constant $\Gamma_c = \Gamma(\zeta_{\text{eff}} \equiv (\delta_{\text{eff}}/l)(j_0/I))$. Noteworthy, Eq. (81) is a universal asymptotic law at current densities $j_0 \geq 2I$, valid for all values g. The current–voltage dependence in this regime, thus, involves a logarithmic term with a double Tafel slope and an ohmic term. The latter contribution becomes dominant for case (3) ($g \gg 1$), which is characterized by prevailing Ohmic losses in the inactive sublayer ($0 < \chi < (1 - \delta_{\text{eff}}/l)$).

Whereas it is intuitively clear that there must be oxygen depletion effects at large current, ad hoc imposition of them is not a safe procedure. For instance Ref. 115, which rests formally on a model with a layout similar to those presented in Refs. 8, 17, presupposes the existence of a limiting current density due to O_2-diffusion limitations, assuming at the same time a possibility of unlimited oxygen consumption. To support the high oxygen consumption rate, corresponding to a large fuel cell current density j_0, the incoming O_2 flux has to be high. As we have explained, this is achieved by the shrinking of the effective thickness δ_{eff}, at the expense of a smaller reaction-active area and increased ohmic losses in the inactive parts of the layer. But an a priori definition of a limiting current density could be misleading. Oxygen depletion in the inactive parts, taken alone, does not necessarily evoke critical or limiting effects in the current–voltage relations. Such an effect, that is, a strong potential drop across the cathode at a certain current density, arises only if the oxygen depletion caused by slow gas diffusion is accompanied by likewise slow proton transport. The cooperation of both mass transport limitations will cause a large potential drop in the inactive parts of the layer. The concerted nature of this effect is illustrated below in Figs. 16 and 17.

We see that a drastic increase in voltage losses appears at a certain current density, but only for the case of poor proton conductivity.

8.2.3.5.3 **Optimum Thickness** The thickness dependence of the performance characteristics can be analyzed generally in terms of the scaling properties of Eqs. (54–56) with respect to l, cf. Sect. 8.2.3.4.1. Consider a catalyst layer of the thickness l at a reference current density j_0. Equation (59) shows that the adjustment to a new thickness $\tilde{l} = \kappa l$ at a simultaneous change of $\tilde{j_0} = j_0/\kappa$ results in a potential shift by $\Delta\eta = -2b\ln\kappa$. If the performance in the original system with thickness l were given by $\eta_0(j_0) = 2b\ln(j_0/\sqrt{Ii^*})$, the rescaled system (indicated by $\sim$) will exhibit the following performance

$$\begin{aligned}\tilde{\eta}_0(\tilde{j_0}) &= 2b\ln\left(\frac{j_0}{\sqrt{Ii^*}}\right) - 2b\ln\kappa \\ &= 2b\ln\left(\frac{j_0}{\kappa\sqrt{Ii^*}}\right) \\ &= 2b\ln\left(\frac{\tilde{j_0}}{\sqrt{Ii^*}}\right) = \eta_0(\tilde{j_0}). \quad (82)\end{aligned}$$

Thus, the electrode potentials η_0 and $\tilde{\eta}_0$ are identical functions of the current density in the original system (thickness l) and in the rescaled system (thickness $\tilde{l} = \kappa l$).

If the performance is determined by $\eta_0(j_0) = b\ln(j_0/i^*)$ the rescaling procedure will give

$$\tilde{\eta}_0(\tilde{j_0}) = \eta_0(\tilde{j_0}) - b\ln\kappa. \quad (83)$$

Thus, the electrode potential in the rescaled system is given by the one in the original system plus a constant shift by $-b\ln\kappa$.

In other words, in the kinetic regime with a predominant simple Tafel dependence a change in l by a factor of κ shifts polarization losses by $-b\ln\kappa$. i^* is the most important parameter here, and a larger active surface due to increased thickness improves the performance.

In the intermediate regime the dominating contribution to η_0 is a $2b\ln(j_0/\sqrt{Ii^*})$ term, and, thus, as demonstrated in general terms, the electrode potential is almost l independent. The catalyst layer thickness could be varied without having to pay a penalty in electrode potential.

In the oxygen depletion regime, an ohmic contribution to electrode potential appears in addition to the $2b\ln(j_0/\sqrt{Ii^*})$ term. This contribution could be dominant for small σ. The effect of thickness in this regime is given by

$$\tilde{\eta}_0(\tilde{j_0}) = \eta_0(\tilde{j_0}) + (1+\kappa)\tilde{l}\frac{\tilde{j_0}}{\sigma}. \quad (84)$$

Evidently, the best compromise between performance losses incurred by the reaction kinetics on the one hand and transport limitations on the other hand, is realized in the intermediate regime. Hence, for a fixed composition and specified target current density j_0 (or j_0 interval) there is an optimum thickness l_{opt}. The same layer may be too thin in the kinetic regime, whereas it may be too thick, when operated in the oxygen depletion regime.

This thickness paradigm, just explained, is visualized in Fig. 14. For a definite electrode composition the optimum current density range (intermediate regime) can be identified by checking the correspondence of the polarization curve to the double logarithmic law $2b\ln\left(j_0/\sqrt{Ii^*}\right)$. Once the limits on j_0 for which this correspondence holds are known, one can draw the phase diagram by using the fact that

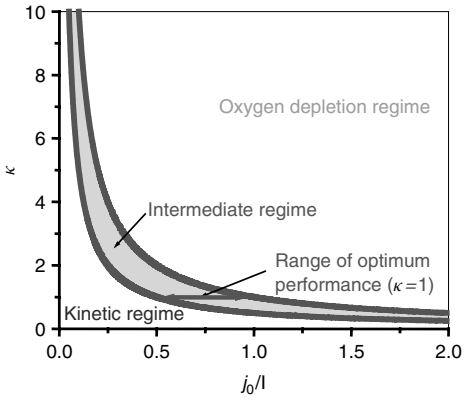


Fig. 14 Phase diagram, which indicates the thickness that a catalyst layer of certain composition should have in order to perform in the intermediate regime (the region between the solid lines, which corresponds to minimal overvoltage losses). κ is the factor of thickness rescaling relative to the reference thickness ℓ ($\tilde{\ell} = \kappa\ell$). In the kinetic regime voltage losses could be reduced by increasing ℓ. In the oxygen depletion regime reduction of ℓ will reduce mass transport losses and thereby improve performance.

these limits vary inversely proportional to the electrode thickness, subject to the invariance properties of Eqs. (54–56). For instance, at the reference thickness l the electrode with the phase diagram depicted in Fig. 14 performs best in the current density range $0.5 < (j_0/I) < 1.0$, where $I = 4F\bar{P}_{O_2}\mathcal{D}/RTl$. A smaller current will ask for a larger optimum thickness, as evident from Fig. 14.

Consider, for example, a catalyst layer with thickness $l = 10\,\mu m$ and $I = 1\ A\,cm^{-2}$ (corresponding to $D \approx 10^{-4}\ cm^2 s^{-1}$ at $T = 300\ K \bar{P}_{O_2} = 1$ atm), which obeys the phase diagram in Fig. 14 (at defined composition). The optimum current density range of operation for this electrode would be $0.5–1.0\ A\,cm^{-2}$. If, however, for example, for reasons of maximum efficiency, the target current density of fuel cell operation is $0.25\ A\,cm^{-2}$ then the electrode thickness should be increased to at least $l = 30\,\mu m$. This would mean that (since composition is kept fixed) the material costs of the catalyst layer would increase as well by a factor of 3. It is thus obvious that a phase diagram of the catalyst layer is a powerful tool for a rational choice of the optimal thickness of the layer.

8.2.3.6 Composition Dependence

In Sect. 8.2.3.5 we have discussed how for the given parameters that characterize ionic transport, gas transport, and reactivity of the catalyst layer, and a specified target current density, one can choose the layer thickness, which provides minimal voltage losses. The parameters, however, depend on composition, which can be varied to optimize the performance. By composition one may imply the chemical composition of the components. Such variation is a subject of material chemistry. Here we will discuss only the variation of the relative amounts of the distinct components (carbon, Pt, PFSI, PTFE) which is a subject of physics of composites.

In spite of the widely recognized importance of an advanced catalyst layer design, detailed structural data for catalyst layers are still scarce in the open literature on fuel cells [116, 117]. In one of the rare experimental studies, Uchida et al. showed the effect of the variation of the PFSI (and PTFE) content on catalyst layer performance [101]. An attempt to rationalize the experimentally observed composition dependence theoretically was first undertaken in Ref. 17. The prerequisites for an adequate theoretical study

are appropriate parameterizations of proton conductivity, oxygen diffusivity, and electrochemically active surface area as functions of the composition. Compared to the initial parameterizations suggested, upgraded versions will be discussed here. The parameterizations in Ref. 17 were based on the two-phase balance of electrolyte and gas pores within the preformed void space left open by the carbon/catalyst phase. In contrast to that, the parameterizations considered here are based on the full balance between varying volume portions of all three components (electrolyte portions, carbon/catalyst phase, and pores).

8.2.3.6.1 **Percolation Properties** The problem of the distribution of the catalyst layer components refers to the class of percolation problems, cf. [91, 118]. Percolating pathways must span through the carbon/catalyst phase, open pore space, and electrolyte network in order to provide an adequate supply of all the reactants and removal of the products. The physics of percolation of GDEs has quite a long history (see Ref. 105 and the literature quoted therein). We will consider here much simpler, ''integral'' concepts, which are quite lessening for the first acquaintance. We also restrict our discussion to the case that the electronically conductive (carbon/catalyst) phase lies above the corresponding percolation threshold. It thus provides a high electronic conductivity and corresponding ohmic losses are negligible.

The percolation threshold is denoted as X_c. It depends on structural properties of the composite. For densely packed composite structures one finds $X_c \approx 0.19$ (for the site percolation) [91]. Weakly interconnected random packings or pore spaces like those resembling a regular diamond lattice have larger values of X_c. The volume portions of the different components are denoted as X_m (metal phase) and X_{el} (electrolyte). The volume portion of the remaining pore space is thus given by $(1 - X_m - X_{el})$. The values of percolation threshold lower than 1/3 give an option of percolation through each of the three components. As it was recently shown [119], the particle size distribution can dramatically diminish the percolation threshold, because smaller particles build bridges in the voids among larger particles of the same sort, which increases connectivity of the network of the given sort of particles.

The law of diminishing is, roughly, given by equation

$$X_c^{poly} = X_c^{mono}/(1 + \Delta/4) \tag{85}$$

Here X_c^{poly} and X_c^{mono} are the percolation thresholds in, respectively, a polydisperse and a reference monodisperse system, whose particle size is equal to the mean particle size in the polydisperse system. Δ characterizes dispersion,

$$\Delta = \frac{\langle r^4 \rangle}{\langle r^2 \rangle^2} - 1 \tag{86}$$

($< r^2 >$ is the mean square radius of the particles), and Eq. (85) is valid for moderate dispersion, $\Delta < 4$. For more details, the reader is referred to Ref. 119, but its main message for us here is that the percolation threshold can easily become as small as 0.1 in dense polydisperse composites or assemblies of soft particles.

Based on these definitions the proton conductivity of the layer is often parameterized as

$$\sigma(X_{el}) = \sigma_s \frac{(X_{el} - X_c)^\tau}{(1 - X_c)^\tau} \Theta(X_{el} - X_c) \tag{87}$$

where σ_s is the bulk-electrolyte conductivity and τ is the critical exponent. Θ is the Heaviside step function, which accounts for zero conductivity below the percolation threshold. Near the percolation threshold critical exponents are universal, that is, independent of the structure of the composite material, but they depend on the dimensionality of the system. In three dimensions $\tau \approx 2$ [91, 118]. Strictly speaking this expression and its universality hold only close to the percolation threshold, but in 3d densely packed lattices it appears to work, empirically, almost in the full range of concentrations, except for the close proximity of $X_{el} = 1$. The latter is, however, not an interesting region for us, because full occupation of the composite by one component will ruin the connectivity in the networks of the other two components, and no one should fabricate such catalyst layers.

Less straightforward than the parameterization of conductivity is the one of oxygen diffusivity. The main pathways of O_2 supply run through the open pore space, in which they exhibit percolation-type dependence. However, there is also a residual O_2 diffusivity through the polymer material, and through nanopores between the carbon/catalyst grains, which are impenetrable for the polymer material. An appropriate parameterization of the diffusivity parameter, introduced in Eq. (42) is, therefore, given by

$$I(X_{el}) = I_s\,[(1-\chi_d)\theta(1-X_{el}-X_m-X_c) \times \left(\frac{1-X_{el}-X_m-X_c}{1-X_c}\right)^{\nu} + \chi_d\Bigg] \quad (88)$$

I_s is the diffusion parameter in the absence of any network constraints, say the one at very high porosity. The exponent ν here can not be exactly the same as in Eq. (78), but we will take the same value for simplicity. The parameter χ_d accounts for the residual gas transport through short "necks" of membrane material. At high residual diffusivity, that is, $\chi_d \approx 1$, the diffusion coefficients would not depend on composition. There are a number of purely empirical approximations for gas diffusivity of the porous layers [80, 88], but we will use Eq. (88) for qualitative estimates.

For the exchange current density it is logical to adopt a parameterization, which rests on the theory of active bonds in dual porous composite materials [120]

$$i^* = i_s^* P(X_{el}) P(X_m)\{(1-\chi_{ec}) \times (1-[1-P(1-X_{el}-X_m)]^M) + \chi_{ec}[1-P(1-X_{el}-X_m)]^M\} \quad (89)$$

Here, i_s^* is the total exchange current density (in A cm^{-2}), proportional to the optimum surface area of all catalyst particles at densest packing of carbon/catalyst particles. It thus, involves, as a factor, the maximum real-to-apparent surface area ratio with typical values 200–400 [121]. The real exchange current density, i^*, involves the reduced real-to-apparent surface area ratio, corresponding to the electroactive fraction of the catalyst surface at random composition, which is obtained subject to the following criteria. The factor $P(X_{el})P(X_m)$ in Eq. (89) takes into account the probability to obtain a meeting point between a carbon/catalyst and an electrolyte particle with both of them connected to their corresponding infinite clusters. $P(X)$ is the density of an infinite cluster of a percolating component, for which one may use a fitting approximation [120],

$$P(X) = \frac{X}{[1+\exp(-a(X-X_c))]^b} \quad (90)$$

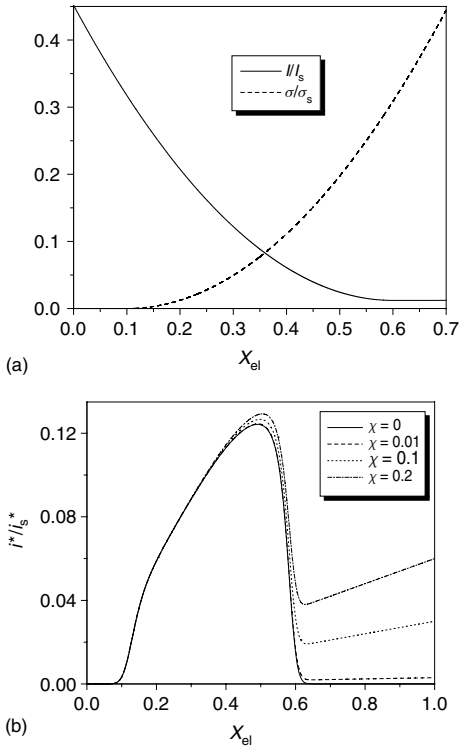


Fig. 15 Dependence of basic catalyst layer parameters on composition, according to Eqs. (87–89). The percolation-type dependencies on electrolyte content X_{el} are depicted for the normalized parameters of (a) proton conductivity, σ, and oxygen diffusivity, I, as well as (b) for the exchange current density at the indicated values of χ_{ec} (residual reactivity), the percolation threshold $X_c = 0.1$, and the residual diffusivity parameter $\chi_d = 0.01$.

With $a = 53.7$ and $b = 3.2$, Eq. (90) reproduces well experimental and simulation data for face centered cubic (fcc) lattice (with no claim on reproduction of critical exponent at the percolation threshold); for other structures the parameters could be different. The first term in curly brackets in Eq. (89) takes into account the probability that the two particles have at least one pore as their bond neighbor that belongs to the infinite pore network, cf. [120]. The parameter M is the average number of bond neighbors (for fcc lattice $M = 4$).

Bonds that are not connected to the infinite cluster of pores can still exhibit a residual electroactivity since they are not entirely disconnected from the oxygen supply. Indeed, usually, there is a continuous pathway through small voids between even densely packed particles; furthermore, oxygen is able to pass the short tracks without voids through the PEM component. This is taken into account by the second term in the curly brackets of Eq. (89). The parameter χ_{ec} accounts for this opportunity, provided by the effect of residual diffusivity of oxygen. A large residual diffusivity ($\chi_d \approx 1$) will make the requirement of accessibility to the infinite gas pore cluster redundant, thus giving $\chi_{ec} = 1$. For small residual diffusivities ($\chi_d \approx 0$), χ_{ec} will also be small, $\chi_{ec} \approx 0$, since reactions will predominantly take place at the fraction of electroactive meeting points that fulfill the requirement of connectivity to the infinite gas pore cluster. The relative importance of gas phase diffusion and accessible three-phase boundary will also depend on electrode thickness. With decreasing thickness, mass transport limitations will be less important, corresponding to larger values of χ_{ec}. For thickness of about 100 nm, $\chi_{ec} \approx 1$.

The inclusion of the residual diffusivity via the "switch parameter" χ_{ec} (which may be called a *coefficient of residual electroactivity*) in Eq. (89) is an interpolation formula. A rigorous treatment would have considered the two contributions to current generation separately, namely, solving the corresponding separate sets of Eqs. (54–56). The first set should involve

gas pore diffusivity to the three-phase boundary of the three percolating components. The second set should explore only the residual diffusivity to the two-phase boundary between carbon/catalyst and electrolyte components. The two contributions to the current generation should be then properly summed up. However, as a mean field approximation, the interpolation scheme, implicit in Eq. (89), is good enough for qualitative estimates, at least outside the region of close vicinity of the point $X_{el} \approx 1 - X_m - X_c$.

Dependencies of the three main parameters on the electrolyte volume portion X_{el} are depicted in Fig. 15. Here, a percolation threshold $X_c = 0.1$ is taken, implying a highly interconnected composite structure. The carbon/catalyst volume portion is kept fixed at $X_m = 0.3$, a value that lies well above the percolation threshold. Consistent with the small percolation threshold a large value $M = 8–10$ was presumed. The residual diffusivity coefficient was taken ad hoc as $\chi_d = 0.01$. The effect of different values of the residual electroactivity parameter χ_{ec} is shown in the plot of i^* in Fig. 15(b).

The plot of i^* as a function of X_e exhibits a maximum in the electroactive fraction $i^*/(i_s^* X_m)$ of catalyst particles (normalized to the metal content). For the specified percolation parameters of the layer, the catalyst utilization corresponding to this maximum is 42%, as obtained at $X_{el} = 0.5$ (and $X_m = 0.3$). This example illustrates the limitations imposed on the catalyst utilization by the percolation properties of the interpenetrating networks.

8.2.3.6.2 Composition-dependent Performance

Polarization curves of catalyst layers at varying compositions are shown in Fig. 16, using composition dependencies as depicted in Fig. 15 and $i_s^* = 10^{-7}$ A cm^{-2}, $I_s = 1$ A cm^{-2}, $b = 30$ mV ($n_t = 1$, $\alpha_c = 1$), $\sigma_s b = 1$ A cm^{-2},

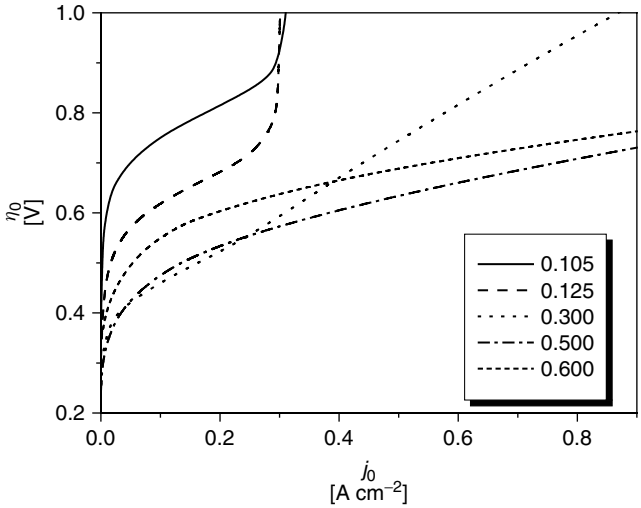


Fig. 16 Catalyst layer polarization relations at various compositions, that is, various electrolyte volume portions X_{el}, indicated in the box, at $I_s = \sigma_s b = 1$ A cm^{-2}, $i_s^* = 10^{-7}$ A cm^{-2}, $b = 30$ mV.

$\chi_d = 0.01$, and $\chi_{ec} = 0.01$. The curves demonstrate that upon increasing j_0 better performance will be obtained with larger electrolyte content. The larger electrolyte volume portions reduce ohmic losses in the layer and thereby improve the performance at large current densities. For increased ratio $I_s/\sigma_s b$ the best performance will be achieved with larger X_{el}, since proton conductivity becomes the rate-determining process. If $I_s/\sigma_s b$ becomes large enough so that sufficient oxygen supply is warranted independent of the percolation properties, the optimum performance will be obtained for $X_{el} = 1$, i.e. we do not need a dual porous composite. This, of course, does not correspond to the real situation, unless ultralow thicknesses ($l \simeq 100$ nm) are considered. For small residual diffusivity, that is, small χ_d, the optimum performance will be obtained at smaller X_{el}.

In Fig. 17 the calculated composition-dependent current–voltage plots of the CCL are compared with experimental fuel cell data by Uchida et al. [101] The parameters used in the model are $i_s^* = 10^{-7}$ A cm^{-2}, $I_s = 1.2$ A cm^{-2}, $\sigma_s b = 1.2$ A cm^{-2}, $b = 30$ mV and $\chi_d = 0.01$ and the fuel cell voltage is given by

$$U_{fc} = 1.23\ V - \eta_0(j_0) - 0.1\frac{\text{cm}^2}{S} j_0 \quad (91)$$

where the open-circuit potential is ≈1.23 V under ambient conditions. The last term on the rhs accounts for overvoltage losses incurred by the membrane (Flemion®, 100 µm thick, conductivity 0.1 S cm^{-1}, no dehumidification, cf. Sect. 8.2.2). Anode losses are neglected. The relation between electrolyte weight content μ_{PFSI} (in mg cm^{-2}) and volume portion, X_{el}, is approximated by $X_{el} = (0.5\ \text{mg}^{-1}\ \text{cm}^2)\mu_{PFSI}$.

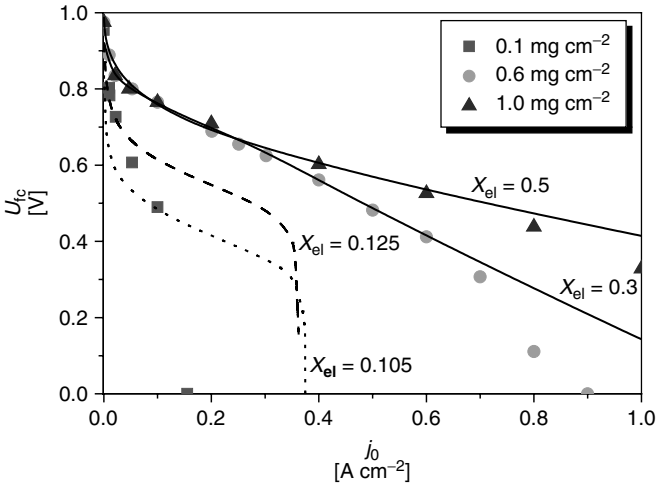


Fig. 17 Comparison of the model calculations with experimental data by Uchida et al.[101, 102]. Current–voltage plots for the fuel cell (including voltage losses due to membrane and cathode catalyst layer). The fits are obtained with the percolation properties of Fig. 15, and $I_s = \sigma_s b = 1.2$ A cm^{-2}, $i_s^* = 10^{-7}$ A cm^{-2}, $b = 30$ mV. A linear relation between PFSI content μ_{PFSI} and X_{el} was used, $X_{el} = (0.5\ \text{mg}^{-1}\ \text{cm}^2)\mu_{PFSI}$.

We see that the main experimentally observed trends are reproduced within the percolation approach. At $j_0 > 0.2\ \mathrm{A\,cm^{-2}}$ the optimal electrolyte content is $1.0\ \mathrm{mg\,cm^{-2}}$. At small current densities, electrodes with reduced electrolyte content can give slightly better performance. For the smallest electrolyte content of $0.1\ \mathrm{mg\,cm^{-2}}$, insufficient proton conductivity leads to electrode failure at $j_0 \approx 0.2\ \mathrm{A\,cm^{-2}}$. The deviation between experimental data and calculated curves at $j_0 > 0.6\ \mathrm{A\,cm^{-2}}$ is probably due to deficient water management in the experimentally studied cell, which leads to electrode flooding, a phenomenon that is not considered in the current model.

8.2.3.6.3 **Optimum Composition** Theory predicts that the optimum current–voltage performance is obtained with ~33 vol% of ionomer, as validated by experiments [122, 123]. The optimum catalyst utilization is, however, found with ~50 vol% of ionomer, as can be seen in Fig. 15. This discrepancy is due to nonoptimal gas diffusion. Facilitating oxygen diffusion would bring the two optima together. As mentioned before, a result of the statistical theory is that the maximum catalyst utilization does not exceed ~40%.

The theory was also used to explore novel design ideas. It was predicted that functionally graded layers with enhanced ionomer content near the membrane interface and correspondingly reduced ionomer content near the GDL side would result in better performances compared to standard CCLs with uniform composition [122]. This prediction was recently verified in experiment [123]. Correspondingly, fabricated catalyst layers with a three-sublayer structure result in enhancements of fuel cell voltage by ~5–10%.

8.2.3.6.4 **Effectiveness Factor of Single Agglomerates** Macrohomogeneous electrode theory, described so far, has been successfully explored in fuel cell diagnostics and optimization [17, 122–126]. Nowadays, finer details of structure and electrocatalytic mechanisms in CLs and model nanoparticle electrocatalysts are moving into focus [127].

The three-phase boundary, prescribing exclusive locations at which reactions proceed, is well defined in regular composite media, as encountered more genuinely in SOFC anodes [120, 128]. In CCLs of PEFCs it cannot be defined precisely, as indicated in Fig. 11(b). Here, catalyst particles acquire individual activities depending on their immediate environment.

It is rather useful to consider agglomerates as structural units of CCLs, cf. Fig. 11(b). Ionomer molecules in micropores could act as binder. Bulky ionomer on the agglomerate surfaces is the proton conductor. Ideal locations of catalyst particles are at the true three-phase boundary between gas pores, bulky ionomer, and Pt/C phase (highlighted by the red star in Fig. 11(b)). How active are catalyst particles at some distance away from this ideal spot, for instance, at the three points indicated by yellow stars in Fig. 11(b)? For each of these points, the contact to gas phase, proton-conducting electrolyte phase or both of these, are impaired. A recent model of performance at the single agglomerate level revealed that the effectiveness of catalyst utilization depends strongly on composition and size of agglomerates and the local operation conditions (electrode potential, concentrations) [129]. Distributions of reactant concentrations, protons, electrode potential, and resulting reaction rates are obtained from solving diffusion and Poisson–Nernst–Planck equations with the Butler–Volmer equation as

a sink term. Corresponding effectiveness factors were found to vary between 30 and 100%. It is vital that micropores inside agglomerates are filled with liquid water in order to keep Pt particles active. What is interesting in this model is the ambivalent role of the potential φ_{el} in the liquid water phase. Protons diffuse through water-filled pores from the surface of the agglomerate toward the center. Under stationary conditions, a gradient φ_{el} is established in order to counterbalance the diffusive proton flux by migration. Therefor, φ_{el} increases from the agglomerate surface toward the center. At the same time, φ_{el} is the driving force for the electrochemical reaction. The variation of φ_{el} could, thus, be rather beneficial in view of achieving uniform distributions of reaction rates.

8.2.3.7 Complex Impedance

An important supplementary tool for performance analysis of PEFC electrodes is the study of the complex impedance, as it provides a tool to monitor changes of electrode function upon variation of its composition. It can help to detect in real time the structural changes due to spontaneous or current-induced repartitioning of the elements of the porous dual percolation network, that could lead to phase segregation and catalyst layer degradation.

The power of electrochemical impedance spectroscopy (EIS) is well acknowledged and widely used in various systems [130, 131]. The history of impedance of porous, distributed electrodes in liquid electrolytes goes back to the early 1960s (see Refs. 132, 133), and some later works [134, 135]. In the context of accumulators, interpretation of electrode impedance was already long ago considered as a promising way to acquire information on the current state of the electrodes [136]. The same refers to the study of corrosion [137, 138]. Nowadays, regular conferences on EIS cover multifarious realms of applications [139, 140].

The potential benefit of impedance studies of porous GDEs for fuel applications has been stressed in Refs. 141, 142. A detailed combined experimental and theoretical investigation of the impedance response of PEFC was reported in Ref. 143. Going beyond these earlier approaches, which were based entirely on numerical solutions, analytical solutions in relevant ranges of parameters have been presented in Ref. 144 which are convenient for the treatment of experimental data. It was shown, in particular, how impedance spectroscopy could be used to determine electrode parameters as functions of the structure and composition. The percolation-type approximations used in Ref. 144, were, however, incomplete, having the same caveats as those used in Ref. 17. Incorporation of the refined percolation-type dependencies, discussed in the previous section, reveals effects due to varying electrode composition and, thus, provides diagnostic tools for optimization of the catalyst layer structure.

8.2.3.7.1 Basic Relations

In order to obtain the complex impedance response of the catalyst layer, the general model, outlined in Sect. 8.2.3.2, has to be amended by the process of double-layer charging in Eq. (45) or (46),

$$\frac{\partial j_p(\chi,t)}{\partial \chi} = -J(\chi,t) - C_{dl}(\eta)\frac{\partial \eta(\chi,t)}{\partial t} \tag{92}$$

Here $J(\chi,t)$ is the rate of current generation which is either given by the RHS of Eq. (45) or (46), and C_{dl} is the double-layer capacitance per unit volume of the distributed electrode. In other words, it is the capacity per unit surface area of the contact between the metal and electrolyte

normalized to the density of such contributions along the coordinate χ. The capacitance per unit surface area could be whatever it should be for that complex interface, including possible constant phase element [130]; the simplest but certainly oversimplified estimate is given by linear Gouy–Chapman theory [145]. Then

$$C_{\mathrm{dl}} = \varepsilon\varepsilon_0 \frac{S_{\mathrm{dl}} l}{\delta_{\mathrm{dl}}}, \tag{93}$$

S_{dl} is the specific area (per unit volume) of the interface between carbon/catalyst and electrolyte. δ_{dl} is the thickness of the double layer of mobile protons and ε is the dielectric constant of the medium in the double layer ($\varepsilon \approx 2$–5). Within the linear Gouy–Chapman model, neither ε, nor δ_{dl} depend on η. If electrostriction effects on the catalyst layer are ignored, S_{dl} neither depends on η, being a constant over the sample. Hereafter, we will adopt the approximation of potential independent C_{dl}. Such a continuous electrical network of the catalyst layer can also be mimicked by a discrete transmission line equivalent circuit. The corresponding Kirchhoff equations of the discrete model are reduced to the above given equations in the continuum limit.

In general, the inclusion of the double-layer charging effects makes it more cumbersome to obtain analytical expressions. In the following only an outline of the solution and a few characteristic expressions will be presented (for details see Ref. 144).

In the stationary problem it is possible to obtain analytical solutions in certain limiting cases, when one of the transport processes causes only negligible losses. These analytical stationary solutions form the basis for the calculation of the dynamic response.

Analytical solutions of the stationary problem for the case of vanishing oxygen transport limitations are given by Eq. (50) in the linear regime of small electrode potentials and by Eqs. (65) and (66) in the Tafel regime. We discuss now the corresponding dynamic response to sinusoidal signals with frequency ω and small amplitude compared to b.

Linear regime: In the small electrode potential regime the complex impedance is

$$Z(\omega) = \frac{1}{\sigma_{\mathrm{p}}} \left(1 + \mathrm{I}\frac{\omega}{\omega_{\mathrm{p}}}\right)^{-1/2} \times \coth\left(\frac{l}{l_{\mathrm{p}}}\left(1 + \mathrm{I}\frac{\omega}{\omega_{\mathrm{p}}}\right)^{1/2}\right) \tag{94}$$

where $\mathrm{I} = \sqrt{-1}$ denotes the imaginary unit and ω_{p} is the characteristic frequency

$$\omega_{\mathrm{p}} = \frac{i^*}{\tilde{b} C_{\mathrm{dl}}} \tag{95}$$

Simple algebraic manipulations of Eq. (94) will give the expression $Z(\omega)$ that was obtained in Ref. 144.

Tafel regime: In order to calculate the dynamic response to a small perturbation $(\delta\eta,\ \delta j)$ at the working point $(\eta_0,\ j_0)$, the complex signals are split into their stationary and dynamic parts in Fourier space,

$$\eta(\chi, \omega) = \eta(\chi) + \delta\eta(\chi, \omega)$$
$$\text{and } j_{\mathrm{p}}(\chi, \omega) = j_{\mathrm{p}}(\chi) + \delta j_{\mathrm{p}}(\chi, \omega) \tag{96}$$

Restriction to linear response requires $\delta\eta/b \ll 1 (b \approx 30$–$50$ mV), resulting in

$$\exp\left(\frac{\eta(\chi) + \delta\eta(\chi, \omega)}{b}\right) \approx \exp\left(\frac{\eta(\chi)}{b}\right)\left(1 + \frac{\delta\eta(\chi, \omega)}{b}\right) \tag{97}$$

The stationary solution at the working point is used in the Fourier transforms of Eqs. (43) and (92). Subtracting the stationary parts from these equations gives

equations on the dynamic response,

$$\frac{\mathrm{d}(\delta\eta)}{\mathrm{d}\chi} = -\frac{1}{\sigma}(\delta j_\mathrm{p}) \tag{98}$$

$$\frac{\mathrm{d}(\delta j_\mathrm{p})}{\mathrm{d}\chi} = -\left(\frac{i^*}{b}\exp\left(\frac{\eta(\chi)}{b}\right) + \mathrm{I}\omega C_\mathrm{dl}\right)(\delta\eta) \tag{99}$$

The substitution of $\eta(\chi)$ from Eq. (65) gives

$$\frac{\mathrm{d}^2(\delta\eta)}{\mathrm{d}\chi^2} = \frac{l^2}{{L_\mathrm{p}}^2(\eta_1)} \times \left(\cos^{-2}\left[\frac{l}{\sqrt{2}L_\mathrm{p}(\eta_1)}(1-\chi)\right] + \mathrm{I}\frac{\omega}{\Omega_\mathrm{p}(\eta_1)}\right)(\delta\eta) \tag{100}$$

The definitions used here are

$$L_\mathrm{p}(\eta_1) = \frac{l_\mathrm{p}}{\sqrt{\alpha_\mathrm{c}}}\exp\left(-\frac{\eta_1}{2b}\right) \tag{101}$$

$$\Omega_\mathrm{p}(\eta_1) = \alpha_\mathrm{c}\omega_\mathrm{p}\exp\left(\frac{\eta_1}{b}\right) \tag{102}$$

A simple analytical solution can be obtained, when the renormalized reaction-penetration depth L_p is large compared to l. Indeed, for this case the variation of $\eta(\chi)$ with χ can be neglected, that is, the electrolyte phase in the electrode is equipotential. Therefore, the dynamic response in the Tafel regime is given by a similar expression as in the linear regime,

$$Z(\omega) = \frac{1}{\Sigma_\mathrm{p}}\left(1+\mathrm{I}\frac{\omega}{\Omega_\mathrm{p}}\right)^{-1/2} \times \coth\left(\frac{l}{L_\mathrm{p}}\left(1+\mathrm{I}\frac{\omega}{\Omega_\mathrm{p}}\right)^{1/2}\right) \tag{103}$$

however, with renormalized penetration depth L_p, (cf. Eq.(101)) characteristic frequency Ω_p, (cf. Eq.(102)) and conductance (cf. Eq. 52)

$$\Sigma_\mathrm{p} = \sqrt{\alpha_\mathrm{c}}\sigma_\mathrm{p}\exp\left(\frac{\eta_1}{2b}\right) \tag{104}$$

Due to the exponential acceleration of the reaction rate with electrode potential, the penetration depth decreases, whereas the characteristic frequency as well as the effective electrode conductivity increase exponentially.

The spatial variation of the stationary electrode potential, $\eta(\chi)$, is small, as long as $L_\mathrm{p} > 2l$. Therefore, the validity criterion of Eq. (103) is

$$2b < \eta_1 < 2b\ln\left(\sqrt{\alpha_\mathrm{c}}\frac{l_\mathrm{p}}{l}\right) \tag{105}$$

If η_1 exceeds this interval, the spatial dependence of $\eta(\chi)$ cannot be neglected. In this case the solution of Eq. (100) leads to a hypergeometric equation, which, due to its complexity, is not shown here.

At $\omega = 0$ the real differential resistance of the layer is recovered (cf. Eq. 51),

$$\Re_\mathrm{diff} = \frac{1}{\Sigma_\mathrm{p}}\coth\left(\frac{l}{L_\mathrm{p}}\right) \tag{106}$$

Several approximations for the calculation of the effect of O_2-transport limitations on the complex impedance response of the electrode have been studied [144]. In one of these approximations, only oxygen diffusion limitations were considered, whereas proton transport limitations were neglected (equipotential electrolyte phase), cf. Sect. 8.2.3.4.4. The stationary solution for this case provides the definition of a gas-diffusion-limited reaction-penetration depth,

$$\frac{\delta_\mathrm{eff}}{l} = \sqrt{\frac{I}{i^*}}\exp\left(-\frac{\eta_0}{2b}\right) \tag{107}$$

as discussed in Sect. 8.2.3.4.4 (cf. Eq. 79).

In the second approximation, both kinds of transport limitations were taken into account in the dynamic equations, but they are considered to be still small enough to ignore spatial dependencies in the stationary parts. Here, in the zero frequency limit the differential resistance is given by similar expressions for $I \gg \sigma b$ (cf. Eq. 51)

$$\Re_{\text{diff}} \approx \frac{b}{\sqrt{\sigma b i^*}} \exp\left(-\frac{\eta_0}{2b}\right) \coth\left[\frac{l}{L_{\text{p}}}\right] \tag{108}$$

and for $I \ll \sigma b$

$$\Re_{\text{diff}} \approx \frac{b}{\sqrt{I i^*}} \exp\left(-\frac{\eta_0}{2b}\right) \coth\left[\frac{l}{\delta_{\text{eff}}}\right] \tag{109}$$

Note that the replacement $\sigma b \leftrightarrow I$ transforms Eq. (101) into Eq. (107) and Eq. (108) into Eq. (109), emphasizing similar impacts of proton and oxygen transport limitations.

The most accurate solution, which does not use any of the aforementioned approximations, can be obtained only numerically; it utilizes the first-order expressions for the spatial variations of the stationary parts of local electrode potential, $\eta(\chi)$, as well as local oxygen partial pressure, $p(\chi)$. These solutions will not be reproduced here. Instead this section will be concluded with a discussion of limiting frequency and thickness regimes, exemplifying the possible use of impedance spectroscopy for catalyst layer diagnosis.

All the approximate analytical expressions that were considered in Ref. 144 give the best results when transport limitations can be considered as insignificant for the stationary electrode performance, corresponding to large renormalized characteristic lengths as compared to the electrode thickness, $\delta_{\text{eff}}, L_{\text{p}} \geq 2l$. These conditions are most likely to be realized at small current densities.

8.2.3.7.2 **Limiting Regimes** *High-frequency limit*, $\omega \gg \Omega_{\text{c}}$. In this regime, the complex impedance has a linear branch at an angle of 45° in the Cole–Cole representation, which is prescribed by

$$Z(\omega) = \sqrt{\frac{1}{\sigma C_{\text{dl}}}}\omega^{-1/2}\frac{\sqrt{2}}{2}(1 - \text{I}). \tag{110}$$

Here, the impedance response is independent of the working point, and the frequency dependence is determined solely by the material parameters of the composite. For $l \ll L_{\text{p}}$, the linear branch appears only at frequencies $\omega > (\sigma/C_{\text{dl}})$. Double-layer charging and proton transport dominate the overall electrode response in this regime, whereas Faradaic processes are insignificant due to the high frequencies. An equivalent representation of this system is an RC-transmission line [130]. Since no fractality or branching of the network is assumed, the response resembles that of a Warburg impedance with a characteristic proportionality $Z \propto \omega^{-1/2}$, where the proportionality constant in Eq. (110) is a characteristic function of the composition. This proportionality constant could be determined in a plot of $|Z|$ versus $\omega^{-1/2}$. Therefore, if σ were known from independent ex situ conductivity measurements or from theoretical considerations based on percolation theory, C_{dl} could be inferred from this kind of plot (or vice versa). Important conclusions may be drawn about the distribution of electrolyte and its effect on performance. A frequency-dependent penetration depth of the sinusoidal signal can be defined by

$$\lambda(\omega) = \sqrt{\frac{\sigma}{\omega C_{\text{dl}}}}l \tag{111}$$

Poor proton conductivity will give a small value of this penetration depth. This will

cause the response given by Eq. (110) to be more pronounced in the impedance spectra.

In the case where only oxygen diffusion transport limitations are considered and proton transport limitations are neglected, the linear impedance response prescribes a perfect semicircle in the complex plane without showing a linear branch in the high-frequency limit. The response given by Eq. (110) is, thus, an exclusive feature of proton transport limitations, which, thereby, provides a feasible tool for their characterization.

Low frequency limit, $\omega \ll \Omega_p$. In this regime, a Taylor expansion in Eq. (103) gives

$$Z(\omega) = \Re_{\text{diff}} \left(1 - A \left(\frac{l}{L_p} \right) \left(\frac{\omega}{\Omega_p} \right)^2 - I \frac{\omega}{\Omega_p} B \left(\frac{l}{L_p} \right) \right) \tag{112}$$

with

$$A(\xi) = \frac{1}{2} \left\{ \frac{3}{4} + 2\xi \frac{3/2 + \xi \coth(\xi)}{\sinh(2\xi)} \right\},$$
$$B(\xi) = \frac{1}{2} \left\{ 1 + \frac{2\xi}{\sinh(2\xi)} \right\}, \quad \xi = l/L_p. \tag{113}$$

This asymptotic expansion forms a part of a semicircle in the complex plane of a Cole–Cole plot. This kind of response indicates charge-transfer limitations, bypassed through the double-layer charging. The center of the semicircle is located on the real axis at

$$M = \Re_{\text{diff}} \left(1 - \frac{B^2}{2A} \right) \tag{114}$$

and its radius is given by

$$R = \Re_{\text{diff}} \frac{B^2}{2A} \tag{115}$$

For $L_p/l \gg 1$ the limitations due to proton transport are practically absent and the impedance response forms a perfect semicircle in the Cole–Cole representation with $M = R = \Re_{\text{diff}}/2$, due to the parallel processes of charge-transfer and double-layer charging, which are distributed homogeneously within the layer. The frequency in the turning point of the semicircle is $\approx \Omega_p$, the approximation being the better the larger the ratio L_p/l. For $L_p/l \cong 1$ the error of this estimate is about 10%.

The linear response to a small perturbation in the Tafel regime is characterized by the renormalized values L_p, Ω_p, and Σ_p. Close to the lower boundary of the potential range, specified in Eq. (105), the penetration depth L_p will be large compared to the electrode thickness. With the increase of the electrode potential, however, L_p decreases exponentially, while Ω_p increases exponentially. Once L_p approaches l, the differential electrode resistance takes the form

$$\Re_{\text{diff}} \approx \frac{1}{\sqrt{\alpha_c \sigma_p}} \exp\left(-\frac{\eta_1}{2b} \right) \tag{116}$$

that is, the Tafel slope doubles if proton transport limitations become as important as kinetic losses.

Thus, the graphical representation of impedance spectra provides definite tools for the determination of catalyst layer parameters via the identification of the linear branch in the high-frequency limit and the branches of semicircular character.

In Fig. 18 composition-dependent impedance spectra are depicted. Catalyst layer properties have been parameterized as functions of the composition as described in Sect. 8.2.3.6.1. For the double-layer capacitance a parameterization in the form.

$$C_{\text{dl}} = C_{\text{dl}}^{\text{s}} P(X_{\text{el}}) P(X_{\text{m}}) \tag{117}$$

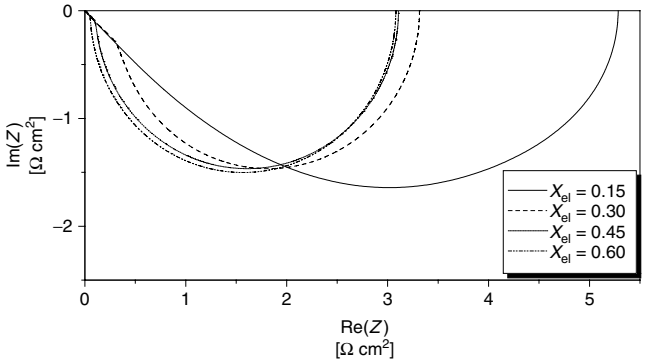


Fig. 18 Composition-dependent impedance spectra at $j_0 = 0.01$ A cm^{-2}. The plotted spectra are calculated from Eq. (103). For high electrolyte contents, $X_{el} > 0.3$, Eq. (92) gives good results. The reference parameters used are $b = 30$ mV, $i_s^* = 10^{-7}$ A cm^{-2}, $\sigma_s = 20$ S cm^{-2}, and $l = 10$ µm. The percolation relations depicted in Fig. 15 control composition dependence.

is used. Here, C_{dl}^s is the double-layer capacitance of a system that is completely filled with electrolyte. The accessibility to the infinite macroporous network is not a prerequisite for a piece of the contact area in order to contribute to the double-layer charging. The impedance spectra were calculated for $j_0 = 0.01$ A cm^{-2} using the approximations that only include proton transport limitations, based on Eq. (100). Since oxygen transport limitations were supposed to be negligible ($I \gg j_0$) in this figure, it was consistent to put $\chi_d = 1$ in Eq. (89). The reference parameters used in this plot are $b = 30$ mV, $i^*/l = 2 \times 10^{-4}$ A cm^{-3}, $\sigma = 20$ S cm^{-2}, and $l = 10$ µm.

The high-frequency domain of the impedance spectra clearly exhibits the linear branch of the response, prescribed by Eq. (110). It is the more pronounced, the more severe proton transport limitations are, that is, the smaller X_{el} is. Calculated parameters corresponding to different compositions are listed in Table 3. At small X_{el} the renormalized penetration depth L_p is smaller than the electrode thickness. Therefore, the stationary potential drop over the catalyst layer is considerable ($\Delta\eta \approx 20$ mV) and Eq. (100) has to be used in order to calculate the impedance response. At larger $X_{el} > 0.30$, one obtains $L_p/l \geq 2$. Hence, the local variations in stationary electrode potential are small, and Eq. (103) provides a good approximation of the complex impedance.

Tab. 3 Calculated parameters, which characterize the catalyst layer performance in the limit of vanishing oxygen diffusion limitations for various compositions at $j_0 = 0.01$ A cm^{-2}. Reference parameters are specified in the caption of Fig. 16

X_{el}	η_l [V]	η_0 [V]	l_p/l	L_p/l	Ω_c s^{-1}
0.15	0.225	0.248	18.3	0.60	902
0.30	0.211	0.214	42.0	1.77	568
0.45	0.198	0.199	58.1	3.04	365
0.60	0.189	0.190	72.0	4.33	277

8.2.3.8 Electrode Characterization and Diagnostics

A simple phenomenological model of the CCL in PEFC helps to rationalize

the impacts of oxygen diffusivity, proton conductivity, and exchange current density on electrode performance. The catalyst layer has a three-phase composite structure, which consists of fractions of carbon/catalyst, gas pores and electrolyte. The performance is calculated and analyzed in a macro-homogeneous approach in which three effective parameters (σ, I, i^*) represent the physical properties of the layer. Although the approach is presented here for the CCL in PEFC, similar approaches can be used to rationalize the operation of anode catalyst layers in DMFC.

The main conclusions are:

Reaction-penetration depths: Characteristic length scales can be defined in various limiting regimes, which represent the penetration of the reaction zone into the layer. In the case of small electrode potentials, the reaction-penetration depth l_p is an important orientation for the optimum thickness with homogeneous catalyst utilization and minimized overvoltage losses. At large electrode potentials, at which reaction rates are described by the Tafel law, the characteristic lengths in the case of dominating proton transport limitations, L_p, or oxygen diffusion limitations, δ_{eff}, decrease exponentially with electrode potential.

Performance regimes: In polarization curves, presented, for example, in Fig. 12, three regimes with dominance of different effects can be discriminated: a kinetic regime at small $j_0 \ll \min(I, \sigma b)$, which exhibits simple Tafel characteristics and homogeneously distributed reaction rates ($l_p \gg l$); an intermediate regime, in which both transport limitations can be considerable and thereby cause inhomogeneous distributions of reaction rates, resulting in a Tafel dependence with a double Tafel slope; and an oxygen depletion regime, in which only a fraction of the layer with thickness δ_{eff} effectively contributes to the current generation with a double Tafel-slope term, whereas the inactive part gives rise to ohmic contributions to electrode potential. Note that it is the concerted effect of proton and oxygen transport limitations that produces the Ohmic corrections. Each of the effects taken alone would give a pure double Tafel-slope dependence in the limit of high j_0.

Optimum thickness. At fixed composition a phase diagram of the catalyst layer can be generated, which establishes a relation between the optimum thickness interval of the catalyst layer and the target current density j_0 (or j_0 interval) of fuel cell operation. The optimum compromise between kinetic losses and mass transport losses is realized in the intermediate regime. The existence of an upper limit on the thickness beyond which the performance would start to deteriorate is due to the concerted impact of oxygen and proton transport limitations, whereas considered separately each of the effects would only serve to define a minimal thickness.

Composition effects: The effect of catalyst layer composition on performance has been studied using percolation-type dependencies. Calculated polarization curves exhibit a good correspondence with experimental data. At a given current density the optimum composition can be determined on the basis of the model calculations. At larger current densities j_0, the optimum composition corresponds to higher electrolyte contents. At first sight, this latter finding may appear counterintuitive since larger electrolyte contents reduce the porosity of the layer, letting oxygen exhaustion become more severe. However, ohmic losses incurred in regions which are partially or completely depleted in oxygen will be smaller due to the enhanced proton

conductivity. The net effect, which can only be seen if the complex interplay of both transport limitations is considered, is an improvement of performance.

Impedance spectroscopy: Features in impedance spectra, like the linear branch (in Cole–Cole representation) due to proton transport limitations in the high-frequency limit or the semicircular character of the response, observed at lower frequencies, give information about the actual morphology of the layer.

Catalyst utilization: Several factors limit the efficiency of catalyst utilization:

1. catalyst particles in nanometers pores (between carbon grains) are impenetrable for electrolyte and, therefore could be utilized only with diminished effectiveness; liquid water saturation of these pores could, however, render them active;
2. of the remaining catalyst particles, only the fraction that is accessible to oxygen and protons through the corresponding percolating networks, as determined by Eq. (89), is electrochemically active; in the example in Sect. 8.2.3.6.1 the maximum fraction amounted to 42%;
3. at large j_0 only a part of the layer with thickness determined by the renormalized reaction-penetration depth due to proton transport (Sects. 8.2.3.3 and 8.2.3.7.1) or by the effective layer thickness due to oxygen diffusion (Sects. 8.2.3.4.4 and 8.2.3.7.1) contributes effectively to the electrochemical current generation.

Advanced design: Control of the catalyst layer thickness and composition provide the means to optimize catalyst utilization and performance issues. Taking together all factors listed above, only 10–20% of the catalyst is effectively utilized in conventional three-phase composite layers. Most of the expensive Pt is not utilized at all or utilized ineffectively. This astounding result is corroborated also by experimental findings. Ostensibly, CCL design is a field where PEFC could be made much better and cheaper.

The routes toward improved design should start at the microscopic level, for example, via site-selective electrochemical deposition of Pt at the active three-phase boundary [146]. At this point, it is hard to estimate extra efforts to be invested in such procedures, but their prospects for optimized catalyst utilization are obvious.

Another alternative that would help to raise catalyst utilization would be to make CLs of extremely thin two-phase composites. Electroactive Pt (eventually deposited on a substrate) should form the electronically conductive phase. The remaining volume should be filled with liquid water. Since the layer is only a two-phase composite, not impregnated with ionomer, the problem of the protonic contact resistance at the PEM|CL interface could be mitigated, making the CCL insensitive to the type of PEM. Using Poisson–Nernst–Planck theory, it could be shown that close to 100% of the catalyst would be utilized, since transport of oxygen and protons would be unproblematic for such thicknesses (~100 nm) [129].

A novel design route expanding on such a two-phase concept has already been followed by the company 3M for some time. The benefits in terms of catalyst utilization and performance enhancement are evident from that work [147]. The activity per total mass of Pt of the 3M layers is larger by about a factor 6 than the mass activity of conventional three-phase CCLs. This affirms the awful inactivity of Pt in conventional CCLs.

Structure–function: at the frontier line: Note that in the case of catalyst layers we deal with very complicated objects. Their structure would not match any ideal lattice model. The particle size distribution notion does not apply straightforwardly because unlike the solid composites for solid oxide fuel cells, the PEM catalyst layer is rather a bicontinuous dispersion than an agglomerate of the grains (particles). On the other hand, they are unlikely volume or surface fractal objects [148] to impose certain generality of the properties due to their fractal geometry [149]. Preparation of catalyst layers as MEAs involves procedures, for example, the hot pressing or spraying techniques [150–152], that still lack sufficient reproducibility, necessary for fundamental studies of the structure–function relationship. Recently developed automated methods of MEA fabrication [153] give a decent level of reproducibility of the stack properties, but not of a single electrode. Furthermore comparison of the structure and performance of different MEAs is often restricted by the proprietary issues that have nothing to do with science. In this situation application of any additional methods of characterization of MEAs (standard porosimetry, SANS, etc.) in an academic environment would be highly valuable.

Using the simple, integral, mean field concepts is ''in accord'' with our poor knowledge of the detailed microscopic structure of MEAs. This may change soon, but probably not before systematic studies driven by these concepts are available.

8.2.4 Quasi-three-dimensional Simulation of a Cell

8.2.4.1 Goals and Methods

The cross section of a fuel cell is shown in Fig. 19. The MEA, which includes backing layers, catalyst layers, and the membrane, is clamped between two metallic or graphite plates with machined channels for feed gases supply. The gases then

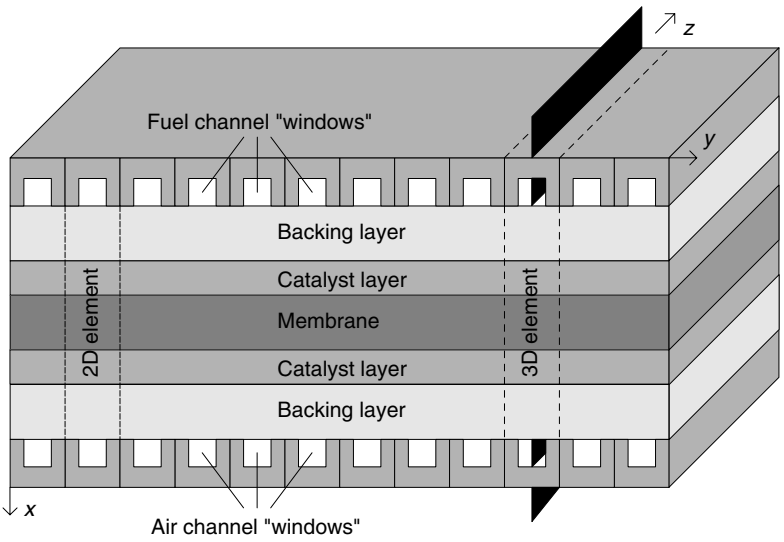


Fig. 19 Sketch of the fuel cell with meander flow field.

pass through the porous backing layers and reach the catalyst layers, where they produce or consume charged particles. Protons produced at the anode move through the membrane to the cathode, while electrons move in the opposite direction through an external load.

Any model of PEFC must cover at least the three basic processes: (1) transport of reactants to/from the catalyst sites, (2) charged particles production or consumption at these sites, and (3) transport of charged particles (electron and proton currents). The simplest realistic model of PEFC must take into account these processes. (More sophisticated models, particularly important for high current regimes, ought to take into account on the same footing the production of water at the cathode side and dynamic liquid–vapor phase balance.)

The first numerical models of PEFC were developed about 15 years ago. Springer et al. published a one-dimensional (1D) steady-state model of PEFC [7]. At the same time Bernardi and Verbrugge developed a model of the PEFC cathode [5] (and later extended it to the whole cell [6]), which in many respects is close to the model of Springer et al.

These models are based on 1D mass and current conservation equations, coupled with Tafel equations in each electrode. Various versions of such models are used nowadays [154–158]. Simplified versions of these models were even amenable to analytical treatment and a number of important aspects of electrode performance were rationalized on the basis of analytical solutions (cf. the preceding section). These models clarified a number of features of PEFC operation and allowed to identify important sources of voltage losses in the cell.

In 1D models it is assumed that gases and potentials are uniformly distributed in the cell lateral cross section. Although this assumption may work in some fragments of the cell, on the scales of the whole cell it is far from real (cf. Fig. 19). There are two main factors that disturb uniformity. First, channels and ribs alternate along the y axis, that is, domains of fixed gas concentration and zero normal current alternate with domains of fixed potential and zero flux of gases. This leads to a complicated 2D field of concentrations and currents in a plane, perpendicular to the channel axis. Second, the feed gas is consumed as it moves along the channel, which leads to nonuniform ''along-the-channel'' distribution of local current density. Furthermore, the two effects are coupled with each other.

More detailed information on cell operation is given by 2D models. These models can be subdivided into two groups. The first group constitute the so-called along-the-channel models, in which equations are written in a plane, directed along the z axis (Fig. 19) [10, 159–166]. Further simplification comes from an assumption that the z-direction components of the fluxes and currents in the MEA are small compared to x-direction components. The 2D problem then is reduced to a set of 1D problems along x direction and to the problem of gas flow in the channel, which supplies the ''boundary conditions'' for the 1D problems [10, 159, 161, 166].

The second group constitute fully 2D ''across-the-channel'' models, which deal with equations in the xy plane [18, 19, 167–171]. These models give 2D fields in the single 2D element (Fig. 19) in the plane perpendicular to the channel, ignoring feed gas consumption along the channel.

Clearly, the complete picture of fuel cell operation would be given by a fully 3D

model. Such a model, however, would be very demanding in terms of computational resources. One of the first 3D models of PEFC was published in Refs. 172, 173. (Historically, the first 3D model of PEFC was developed in Refs. 174, 175. The authors simulated several adjacent 3D elements, shown in Fig. 19, but the length of their computational domain along axis z (Fig. 19) was less than 1 cm.) A small fragment of the structure, shown in Fig. 19 (3D element) which includes a 10-cm fragment of the feed channels was simulated [172, 173]. Presumably due to computational limitations, this model does not resolve catalyst layers, which are represented by effective boundary conditions.

Fortunately, the fuel cell is a two-scale system. The two characteristic length scales are the MEA thickness l_{MEA} and the length of the feed channel L. The latter is 3 to 4 orders of magnitude larger than the former. This allows developing an efficient Q3D model, which combines the advantages of along- and across-the-channel models. This model gives a ''performance map'' in xy cross section of a cell, shown in Fig. 19, taking into account feed gases consumption in the channels. The model makes it possible to investigate the interplay between the small- and large-scale phenomena in fuel cells. Developed for parallel computers, it allows simulating fuel cells with feed channels up to several meters long.

The Q3D model rests on the following idea. Along-the-channel models clearly show [10, 159] that the variation of the local current density and feed gas concentration along the channel are negligible on a length of the order of the MEA thickness. This permits one to neglect the z component of fluxes in mass and current conservation equations, written for the xy plane. In the channel, however, the gas flows mainly along the z axis. Therefore, the fully 3D problem can be split into a 2D problem in the porous layers (*internal problem*) and a problem of gas flow in the channels (*channel problem*). In the simplest variant the variation of the feed gas concentration along the channel can be described by 1D equations. (A more accurate would be 2D or even 3D descriptions of flows in the channel. However, the crucial effects caused by the feed gas exhaustion can be captured, as is shown below, within the scope of a 1D model.)

The outline of calculations is as follows. Let the cell voltage be fixed. For given along-the-channel profiles of feed gas concentrations, the internal problem returns the profile of the local current density. This local current density is then used to update the concentrations of feed gases in the channels. This procedure is repeated until convergence is reached. Thereby a point on the voltage current curve and a map of parameters in a cell cross section is obtained.

The cross section of a fuel cell consists of a number of geometrically identical (or similar) 2D elements (Fig. 19). The internal problem is formulated for a single element. The code is designed for a parallel computer, so that each element ''is solved'' on a separate processor. At the end of each iteration step adjacent elements exchange their ''boundary conditions'' (concentrations and currents along the dashed lines in Fig. 19).

8.2.4.2 Physical Model of a Single 2D Element

8.2.4.2.1 **Basic Assumptions** In order to demonstrate the capabilities of the Q3D model we consider in this chapter a basic

variant of it, which rests on a number of simplifying assumptions. Our basic conjecture will be that the membrane is impermeable to gases and fully hydrated. The assumption of impermeability to gases is justified for PEFC. In liquid feed DMFC methanol permeation through the membrane (crossover) leads to large overvoltage losses. The methanol crossover in gas-fed DMFC (which are considered here) is lower [176]. In this simplified description we will neglect the crossover.

The membrane is assumed to be fully hydrated. The model of membrane water management, discussed in Sect. 8.2.2, suggests that for now, in the mostly used Nafion-type membranes, with thickness in the range of 50 μm, the critical current density of membrane dehydration exceeds by far the typical current densities of fuel cell operation ($<1\ \mathrm{A\,cm^{-2}}$). This inference is further supported by recent experimental studies, which corroborate that the membrane regulates water fluxes in the fuel cell but its own state of hydration is not critically affected by them [125, 126].

Further assumptions are as follows:

1. The cell is isothermal.
2. In the absence of pressure gradients, diffusion due to concentration gradients is the main mechanism of gas transport in the porous layers. Pressure gradients will give an additional contribution to the flux.
3. The local reaction rate in each of the electrodes is described by a variant of the Tafel equation with effective parameters, that are, in first approximation, considered to be constant over each catalyst layer.

The first assumption is usually fulfilled if the cell is under proper thermal management. The second assumption is standard in the theory of fuel cells (for the discussion see Refs. 18, 19). The scope of the third assumption will be discussed below.

8.2.4.2.2 **Model of Gas Flow in Porous Layers** The molar flux of the ith gas component of molar fraction ξ_i is determined by the following relation

$$\mathbf{N}_i = -cD_i\nabla\xi_i + c\xi_i\boldsymbol{v}^{\mathrm{D'Arcy}} \qquad (118)$$

where

$$\boldsymbol{v}^{\mathrm{D'Arcy}} = -\frac{k^p}{\mu}\nabla p \qquad (119)$$

is the D'Arcy velocity of the mixture. The first and second terms on the right-hand side of Eq. (118) describe the diffusion flux due to a concentration gradient and the flux caused by a pressure gradient, respectively. Some of the published results were obtained with a more complicated expression for the flux, in which diffusion in the backing layers was described by Stefan–Maxwell relations [177, 178]. However, the simple diffusion term in Eq. (118) with effective diffusion coefficient D_i (see below) gives practically the same result and essentially simplifies the computational work.

The value of D_i is determined as follows. Under standard atmospheric conditions the mean free path of molecules by the order of magnitude is $l_{\mathrm{free}} \approx 1/(N_{\mathrm{a}}\sigma_{\mathrm{g}})$, where $N_{\mathrm{a}} = 2.68 \times 10^{19}\mathrm{cm}^{-3}$ is the number of molecules per cubic centimeter and $\sigma_{\mathrm{g}} \approx 10^{-14}\mathrm{cm}^2$ is the mean cross section of molecules collision. For pressure in the range 1–3 bar this gives $l_{\mathrm{free}} \approx 10^{-6}\mathrm{cm}$.

Usually, the mean radius $\bar{r}$ of pores in the catalyst layer will be small compared to l_{free} [101, 102]. This means that diffusing molecules will collide more frequently with pore walls than with each other. The prevailing mechanism of gas transport in

the catalyst layer with typical pore sizes in the nm range is, therefore, Knudsen diffusion with diffusion coefficient

$$D_i^K = \psi \bar{r} \sqrt{\frac{8RT}{\pi M_i}} \tag{120}$$

Here the factor ψ accounts for effects of composition (cf. Sect. 8.2.3.6), tortuousity and psd and the square root is the mean thermal velocity of molecules of mass M_i. The value D_i^K serves as a reasonable estimate of the diffusion coefficient.

On the contrary, in the backing layers, pore radius is large as compared to the mean free path and the mechanism of gas transport is free molecular diffusion. This can be characterized by the mean binary diffusion coefficient D_{ij} [103]. The effective diffusion coefficient in backing layers, D_i^B, is determined by

$$\frac{\varsigma}{D_i^B} = \sum_j \frac{\xi_j}{D_{ij}} \tag{121}$$

where ς is the correction factor for porosity.

D_i^K is usually an order of magnitude lower than D_i^B. To avoid infinite gradients in the diffusion coefficient at the catalyst/backing layer interface and also to reflect the continuous character of the diffusion coefficient at the real interface, which is never sharp, D_i is approximated by

$$D_i = D_i^B + (D_i^K - D_i^B)\frac{1}{2}\left(1 \pm \tanh \times \left(\frac{x - \bar{x}}{\Delta}\right)\right) \tag{122}$$

Here $\bar{x}$ is the location of the backing/catalyst layer interface, Δ is the characteristic width of the interface smearing, which must be much smaller than the catalyst layer thickness for this interpolation to work properly (usually Δ is about 1/10th of the catalyst layer thickness); + and − stand for the anode and the cathode respectively ($x = 0$ is at the anode side).

Mass conservation relation for the ith gas component reads

$$\nabla \cdot \mathbf{N}_i = \frac{S_i}{nF} Q \tag{123}$$

where Q is the rate of electrochemical reaction, n is the number of electrons participating in the reaction and S_i is stoichiometry coefficient. Substituting Eqs. (118) and (119) into Eq. (123) we obtain an equation for ξ_i, which contains the pressure gradient.

The equation for pressure p follows from the following arguments. Neglecting the mass of protons we can consider that there are no volume sources and sinks of mass on either side of the cell. The continuity equation for the total mass flux thus reads

$$\nabla \cdot (\rho \boldsymbol{v}^{\mathrm{D'Arcy}}) = 0 \tag{124}$$

where $\rho = cM$ is gas density, M is the mean molecular weight of the mixture, determined by $M = \sum_i \xi_i M_i$. Taking into account Eq. (119) and using the ideal gas law $p = \frac{\rho}{M} RT$ we obtain

$$\nabla \cdot \left(\frac{M}{RT} p \nabla p\right) = 0 \tag{125}$$

Note that the gas composition varies over space; M, therefore, depends on the coordinates.

Since the membrane is assumed to be impermeable to gases, Eqs. (118), (123), and (125) can be written for the anode and the cathode compartments separately with the boundary condition of zero flux at the catalyst layer/membrane interfaces.

8.2.4.2.3 **Potentials** The potential drop across the metal/electrolyte interface is the driving force for the electrochemical reactions. The rates of electrochemical reactions depend on the potential in the membrane electrolyte phase, φ_{el}, and on the potentials of carbon phase φ_a and φ_c on anode and cathode side, respectively (see next section).

These potentials are governed by proton and electron current conservation equations:

$$\nabla \cdot (\sigma_{el} \nabla \varphi_{el}) = Q \tag{126}$$

$$\nabla \cdot (\sigma_a \nabla \varphi_a) = Q_a \tag{127}$$

$$\nabla \cdot (\sigma_c \nabla \varphi_c) = -Q_c \tag{128}$$

where subscripts a and c stand for the anode and cathode respectively, σ_{el} is the proton conductivity of the electrolyte phase, σ_a, σ_c are conductivities of carbon cloth at the anode and the cathode respectively, $Q = -Q_a$ in the anode catalyst layer and $Q = Q_c$ in the CCL. The reaction rates, being finite in catalyst layers, are zero outside. In Sect. 8.2.3 of this chapter we have learned that generally conductivity of membrane phase in the catalyst layer is lower than that of the bulk membrane, since membrane phase never occupies the full space of the catalyst layer. In the simulations, the results of which will be shown below, σ_{el} is put constant through the whole cell for simplicity. In fact, one should have interpolated σ_{el} from its bulk value in the membrane and its value in the catalyst layer. That would have increased the ohmic losses in the MEA and quantitatively affected the performance map (but, presumably, not qualitatively, unless the MEA is poorly fabricated and σ_{el} is orders of magnitude smaller than in the membrane bulk!).

8.2.4.2.4 **Reaction Rates** Electrochemical reactions of methanol or hydrogen oxidation at the anode and oxygen reduction at the cathode are complicated multistage processes. Their studies have a long history [179, 180]. The dependencies of the corresponding reaction rates on the local electrode potential may be determined by a system of equations each reflecting different stages, including adsorption and transformation of intermediates. The resulting dependence may not fit the simple Tafel form, although resembling it in different potential regions.

We will not go into details of a microscopic modeling of these reactions here but rather parameterize the reaction rates by standard approximations. We adopt the Tafel relation, given at the cathode side by

$$Q_c = i_c \left(\frac{c_{O2}}{c_{O2ref}} \right)^{\gamma_c} \exp\left(\frac{\alpha_c F}{RT} \eta_c \right) \tag{129}$$

and at the anode side by

$$Q_a = i_a \left(\frac{c_M}{c_{Mref}} \right)^{\gamma_a} \exp\left(\frac{\alpha_a F}{RT} \eta_a \right) \tag{130}$$

Here, c_{O2} and c_M are oxygen and methanol concentrations in the respective catalyst layer, c_{O2ref} and c_{Mref} are reference concentrations, $\eta_c = \varphi_{el} - \varphi_c$, $\eta_a = \varphi_a - \varphi_{el}$ are the cathode and the anode electrode potentials respectively, i_a, i_c are exchange current densities per unit volume (A cm^{-3}), α_a, α_c are transfer coefficients, and γ_a, γ_c orders of reaction.

As discussed already in Sect. 8.2.3, there is a variety of publications that report experimental evaluations of the parameters α and ν (see e.g., [181–186]). They are often different in various potential ranges, but also depend strongly on the particular reaction environment (the structure and

composition of the catalyst). The parameters, which are used in the simulations, are indicated in the corresponding tables.

All that the models of the class described above need is the ''conversion function'' that defines the relationship between the local values of the reacting gas concentration, overpotential, and the produced current. The subject of the microscopic theory is to justify this relationship.

Sufficiently far from equilibrium, which is the only interesting regime for current generation, the Tafel law (129), (130) with effective parameters is used in the simulations. Measurements performed on model electrocatalysts suggest different values for these parameters depending on the particular catalysts used or just their different structure. The mesoscopic effects in electrocatalysis are the hot topic nowadays [127, 187–189] and they are probable candidates for one of the reasons of this variance. The main cause of this variance may be associated with the complicated interplay between adsorption and reaction stages, and a modeler interested in fuel cell operation should simply accept this variance as a fact. For instance, in Refs. 180, 184, 185 and even in recent measurements for one family of Pt/Ru catalysts (Wipperman and Havranek, unpublished data, see Table 4) the values of the effective Tafel slope $RT/\alpha F$ vary significantly. Importantly, the latter does not remain constant but varies either continuously or stepwise with overvoltage.

All in all a unified approximation (129), (130) with constant coefficients in the whole range of relevant overvoltages is a model simplification. If the variation is in fact dramatic, this will affect the current–voltage plot. Using a unified single law in the whole range is a tool to focus on other effects, such as gas transport or proton transport limitations with a most sketchy account of kinetic reaction limitations. Obviously, if we had predominantly reaction kinetics determining the operation regime, the current–voltage plot would follow the law that we had put in to approximate the ''conversion function''. If the approximation is incomplete we will not have an adequate description of the

Tab. 4 Tafel slopes of the methanol oxidation on technical Pt/Ru anodes. The anodes are prepared by wet spraying. Methanol concentration: 1 M, potential range used for the calculation of the *b* factors: 0.25–0.4 V (RHE) (A. Havranek and K. Wippermann, unpublished results)

Catalyst	*Fractions of catalyst and Nafion [wt%]*	*Pt/Ru-loading [mg[a] cm^{-2}]*	*Press force[a] [kN[a] cm^{-2}]*	*Operating temperature [°C]*	*Tafel slope [mV]*
30% Pt–Ru (1:1) on Vulcan XC 72 (Johnson Matthey)	70/30	1.3	0.5	50	60
40% Pt–Ru (1:1) on Vulcan XC 72 (E-TEK)	70/30	1.0	0.5	50	88
Pt–Ru (1:1) black (Johnson Matthey)	93/7	0.5	0.5	60	56

[a] Press force applied to the samples during the hot-pressing process, hot-pressing temperature: 130 °C.

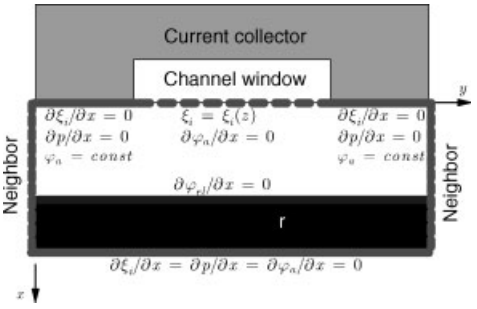


Fig. 20 Boundary conditions for gases, concentration and potentials. The computational domain for the anode side gases and potentials is depicted by bold line (the domain for membrane phase potential φ_{el} includes both catalyst layers and membrane). Shown are conditions at the anode; the same conditions are imposed on the cathode side. Conditions along the left and right sides are obtained on each iteration step from the ''neighbors'' (physical boundary conditions are fixed at the left side of the leftmost and at the right side of the rightmost elements).

electrode current–voltage plot. But this is not a reason for pessimism! The numerical models can be easily extended *to any form of conversion function*, whereas the first results, from which one expects to get the first qualitative insights on the system behavior, should of course start with the simplest approximations for conversion function.

8.2.4.2.5 **Boundary Conditions** The boundary conditions for gas concentration and potentials in a single 2D element (Fig. 19) are shown in Fig. 20. Since the membrane is impermeable to gases, the gradients of all variables (except φ_{el}) along the normal to the catalyst layer/membrane interface are zero. Pressure and gas concentrations along the channel/backing layer interface are determined from the channel problem. Outside the channel the current collectors fix the potentials and form impermeable walls for gases and pressure (where the gradients of these values in x direction are zero). Along left and right surfaces ''boundary conditions'' are obtained from adjacent elements on each iteration step. The other conditions are evident.

8.2.4.2.6 **Finite-difference Approximations** Equations (125–128) are particular cases of the prototype equation

$$\frac{\partial}{\partial x}\left(\Lambda(u)\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(\Lambda(u)\frac{\partial u}{\partial y}\right)=q \tag{131}$$

where $\Lambda(u)$ stands for the corresponding kinetic coefficient. In order to cast this equation into the finite-difference form, we introduce an orthogonal grid in the domain shown in Fig. 20 and approximate Eq. (131) by

$$\frac{0.5(\Lambda^m_{i+1,j}+\Lambda^m_{i,j})\left(\dfrac{u^{m+1}_{i+1,j}-u^{m+1}_{i,j}}{\Delta x_i}\right)-0.5(\Lambda^m_{i,j}+\Lambda^m_{i-1,j})\left(\dfrac{u^{m+1}_{i,j}-u^{m+1}_{i-1,j}}{\Delta x_{i-1}}\right)}{0.5(\Delta x_{i-1}+\Delta x_i)}$$
$$+\frac{0.5(\Lambda^m_{i,j+1}+\Lambda^m_{i,j})\left(\dfrac{u^{m+1}_{i,j+1}-u^{m+1}_{i,j}}{\Delta y_j}\right)-0.5(\Lambda^m_{i,j}+\Lambda^m_{i,j-1})\left(\dfrac{u^{m+1}_{i,j}-u^{m+1}_{i,j-1}}{\Delta y_{j-1}}\right)}{0.5(\Delta y_{j-1}+\Delta y_j)}$$
$$=q^m_{ij} \tag{132}$$

Here, $\Delta x_i=x_{i+1}-x_i$; $\Delta y_j=y_{j+1}-y_j$ are the distances between adjacent

nodes along x- and y axis respectively. This approximation leads to a system of linear equations [18], which can be solved by any standard procedure [190]. In Eq. (132) m is the iteration number. Note that coefficients and the right-hand side of Eq. (132) are evaluated at the previous iteration step.

The prototype of Eq. (123) with the flux given by Eq. (118) is

$$\nabla \cdot (\Lambda \nabla u + \mathbf{W}u) = q \tag{133}$$

Here, the flux is $\mathbf{N} = \Lambda \nabla u + \mathbf{W}u$, where $\Lambda(\mathbf{r})$ and $\mathbf{W}(\mathbf{r})$ are functions of coordinates. (In the whole system of equations these functions in fact depend on the other unknown variables, whose values are taken from the previous iteration step.) This equation is discretized with the Scharfetter–Gummel scheme. (The idea of this scheme was first offered by Allen [191] for hydrodynamic equations and then rediscovered in other fields.) The flux through the ''East'' surface of the computational cell (cf. Fig. 21) has a form

$$N_E^{m+1} = -\frac{W_E^m}{\exp(\beta_E^m) - 1} \times (u_{i+1,j}^{m+1} - \exp(\beta_E^m) u_{i,j}^{m+1}),$$
$$\beta_E^m = \frac{h_i W_E^m}{\Lambda_E} \tag{134}$$

where $\Lambda_E = 0.5(\Lambda_i + \Lambda_{i+1})$ and m enumerates iterations. Similar relations are written for the fluxes N_{W}^{m+1}, N_N^{m+1}, N_{S}^{m+1} through the other surfaces of the computational cell. The equation $\nabla \cdot \mathbf{N} = q$ is then approximated as

$$\frac{N_E^{m+1} - N_{\mathrm{W}}^{m+1}}{0.5(\Delta x_{i-1} + \Delta x_i)} + \frac{N_N^{m+1} - N_{\mathrm{S}}^{m+1}}{0.5(\Delta y_{j-1} + \Delta y_j)} = q_{i,j}^m, \tag{135}$$

which leads to a linear system of equations with respect to $u_{i,j}^{m+1}$.

8.2.4.3 The Channel Problem

The steady gas flow in a long macroscopic channel with impermeable walls is basically a Poiseuille flow with the constant velocity determined by the pressure gradient. However, the velocity of the flow in the fuel cell channel varies since there is mass and momentum transfer through the channel/backing layer interface.

For simplicity we will ignore momentum loss due to meander turns and assume that the channel is straight with the axis directed along z. Consider a cathode channel. Due to the electrochemical reaction each oxygen molecule is replaced with two water molecules in the flow. The continuity equation, therefore, reads (In [192]

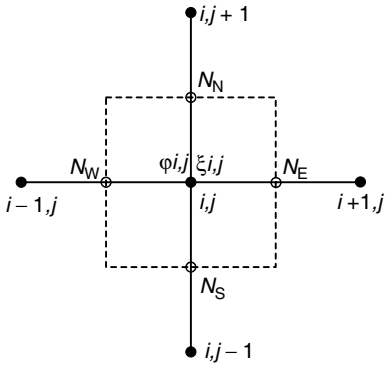


Fig. 21 Sketch of the computational cell (dashed square). Note that concentrations and potentials are given at the nodes, whereas fluxes and currents are given at the surfaces of computational cells.

water flux through the membrane is considered. This flux can essentially affect velocity in the channels. Here all fluxes through the membrane for simplicity are ignored.) [192]:

$$\frac{\partial(\rho_c v_c)}{\partial z} = \frac{j(z)}{h_c}\left(\frac{2M_w - M_{O2}}{4F}\right) \tag{136}$$

where $v_c(z)$ is the flow velocity (in z direction), $\rho_c(z)$ is the flow density, $j(z)$ is the local current density, h_c is the channel height, and M_w and M_{O2} are molecular weights of water and oxygen respectively.

Since the axial component of velocity near the channel/backing layer interface is zero, each consumed oxygen molecule removes momentum from the flow. Each incoming water molecule also removes momentum, since it must be accelerated to the flow velocity. The momentum balance (the Euler equation), therefore, has a form [192]

$$\rho_c v_c \frac{\partial v_c}{\partial z} = -\frac{\partial p_c}{\partial z} - \frac{j v_c}{h_c}\left(\frac{2M_w + M_{O2}}{4F}\right) \tag{137}$$

where $p_c(z)$ is the pressure.

Detailed analysis shows [192] that in the practically interesting range of inlet velocities, being much smaller than the speed of sound, the flow in the channel is incompressible, that is, $\rho_c = const$. Solving Eq. (136) for velocity we immediately obtain

$$v_c(z) = v_c^0 + \frac{1}{\rho_c h_c}\left(\frac{2M_w - M_{O2}}{4F}\right) \times \int_0^z j(z)\,dz \tag{138}$$

The velocity in the cathode channel, therefore, increases. Physically, oxygen molecules are replaced with water molecules in the flow; the total mass of the flow hence increases and to preserve constant density each elementary volume of gas expands while it moves that is, the velocity must increase.

These arguments are valid for the anode channel as well. For instance, the velocity of the flow in the gas-fed DMFC anode channel is given by

$$v_a(z) = v_a^0 - \frac{1}{\rho_a h_a}\left(\frac{M_M + M_w - M_{CO2}}{6F}\right) \times \int_0^z j(z)\,dz \tag{139}$$

where M_M and M_{CO2} are the molecular weights of methanol and CO_2 respectively. Here, the flow loses mass (hydrogen atoms are consumed) and its velocity decreases to preserve the density constant.

Equations (138) and (139) allow to calculate $v_a(z)$ and $v_c(z)$, if $j(z)$ is known (the latter is determined from the solution of the internal problem). Then one can find the feed gas concentration along the channel. For example, in the cathode channel we have

$$\frac{\partial(c_{O2} v_c)}{\partial z} = -\frac{j(z)}{h_c}\frac{1}{4F} \tag{140}$$

where c_{O2} is the oxygen molar concentration. Under given $v_c(z)$ and $j(z)$ integration of Equation (140) is trivial. Analogous algebra determines the methanol or hydrogen concentration in the anode channel. For the case, for example, of methanol we have

$$\frac{\partial(c_M v_a)}{\partial z} = -\frac{j(z)}{h_a}\frac{1}{6F} \tag{141}$$

8.2.4.4 Iteration Procedure

Having described the approaches to the internal and channel problems we now can describe the iteration procedure:

Step 0:

- In all elements put $c_{O2}(x, y, z) = c^0_{O2}$, $c_M(x, y, z) = c^0_M$, $\varphi_c(x, y) = 0$, $\varphi_a(x, y) = \varphi^0_a$, $\varphi_{el}(x, y) = k\varphi^0_a$ (where the constant $k < 1$ is determined from $l_c Q_c = l_a Q_a$; Q_c and Q_a are given by Eqs. 129 and 130).
- Solve Eqs. (123), (125), and (126) to Eq. (128) for each element (Fig. 19)
- Each element sends boundary values (along vertical dashed lines in Fig. 19) to the neighbors and receives their boundary values. The values along the left boundary of the leftmost element and along the right boundary of the rightmost element are determined by physical boundary conditions (zero flux along y axis).
- Calculate the local current density in every element, using the relation

$$j = \frac{1}{l}\int_o^l \mathrm{d}x \int_0^H \mathrm{d}y\; Q_c \qquad (142)$$

where H is the size of the single element along y (cf. Fig. 20), l is the catalyst layer thickness. This gives the local current density in all "windows" (several points along the channel, cf. Fig. 19).

- Interpolate these values linearly. Use the obtained $j(z)$ to calculate along-the-channel profiles of concentrations of oxygen $c_{O2}(z)$ and methanol $c_M(z)$, according to Eqs. (140) and (141) respectively.

Repeat the following procedure until convergence is reached:

- With $j(z)$, $c_M(z)$, $\varphi_c(x, y)$, $\varphi_a(x, y)$, and $\varphi_m(x, y)$ obtained from previous iteration, solve Eqs. (123), (125) and (126–128) for each element.
- Exchange boundary values with the neighbors, as described in step 0.

Calculate local current density with Eq. (142) and determine new $c_{O2}(z)$ and $c_M(z)$.

8.2.4.5 Results

Here we will present several simulation results, which show how the knowledge of the "functional map of the cell" may help to improve the cell performance. Most of the examples presented here refer to gas-fed DMFC, although with minor modifications they can be (and were) applied to hydrogen PEFC. We end this section with a special discussion of the features of the hydrogen cell.

8.2.4.5.1 Gas-feed DMFC (Single Element): Embedded Current Collectors Figure 22 displays the distribution of current densities and reaction rates (the mean current density in the cell is 0.2 A cm^{-2}, the operation conditions are given in Ref. 19) The maps of electron current density in both the anode and the cathode reveal strong peaks near the edges of current collectors. These edges collect all the current produced in front of the gas channels. The peak current density at the edges is about 10 times higher than the mean current density through the cell and it may produce local overheating.

The reaction rates Q_a and Q_c are reduced by the lack of methanol/oxygen in the parts of the catalyst layers "shaded" by current collectors. It is possible to remove the precious catalyst from the shaded regions practically without degradation of performance [18]. However, it is much better to change the design of current collectors. To prevent partial "shading" of catalyst layer by current collectors, it is

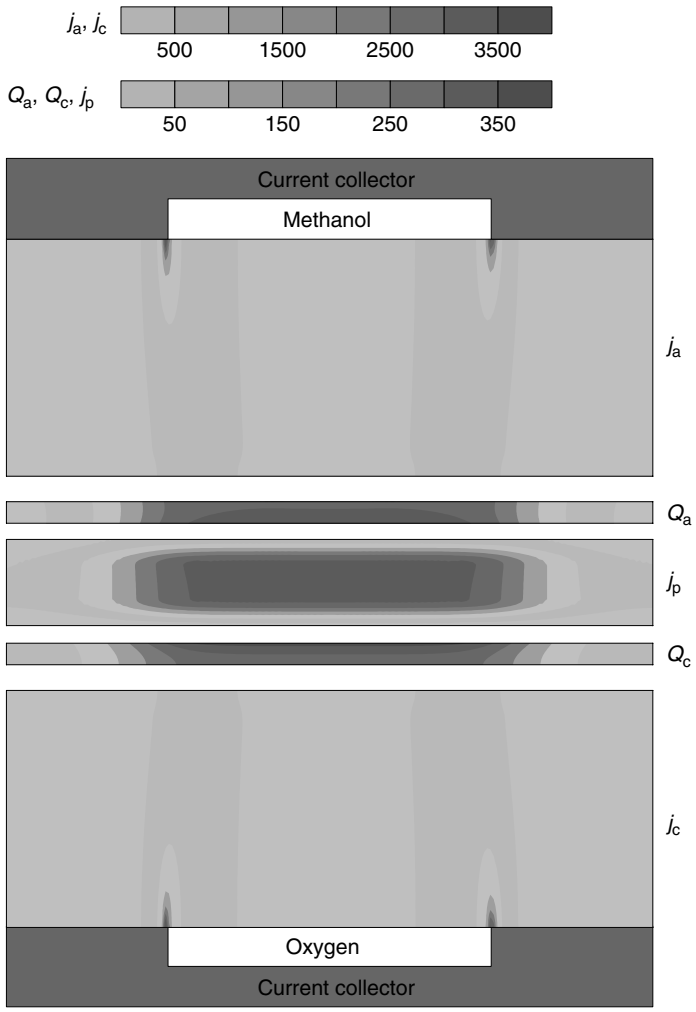


Fig. 22 Cell element with conventional feeding through the channels. Shown are maps of electron current densities j_a, j_c (mA cm^{-2}), reaction rates Q_a, Q_c (A cm^{-3}), and proton current density j_p (mA cm^{-2}).

beneficial to embed current collectors into backing and catalyst layers so that the collector is parallel to the fuel flow (Fig. 23).

The maps of current densities and reaction rates for the embedded design are shown in Fig. 23 (at the same mean current density through the cell of 0.2 A cm^{-2}). It is seen that the distribution of current densities along the surface of the current collectors on both sides of the cell is almost uniform. Reaction rates are also distributed almost uniformly along the catalyst layers, due to uniform (along y) supply of feed gases. The advantage of this geometry is evident. Note that the flow of electrons in that case is perpendicular to the flow of protons. It is, therefore, not

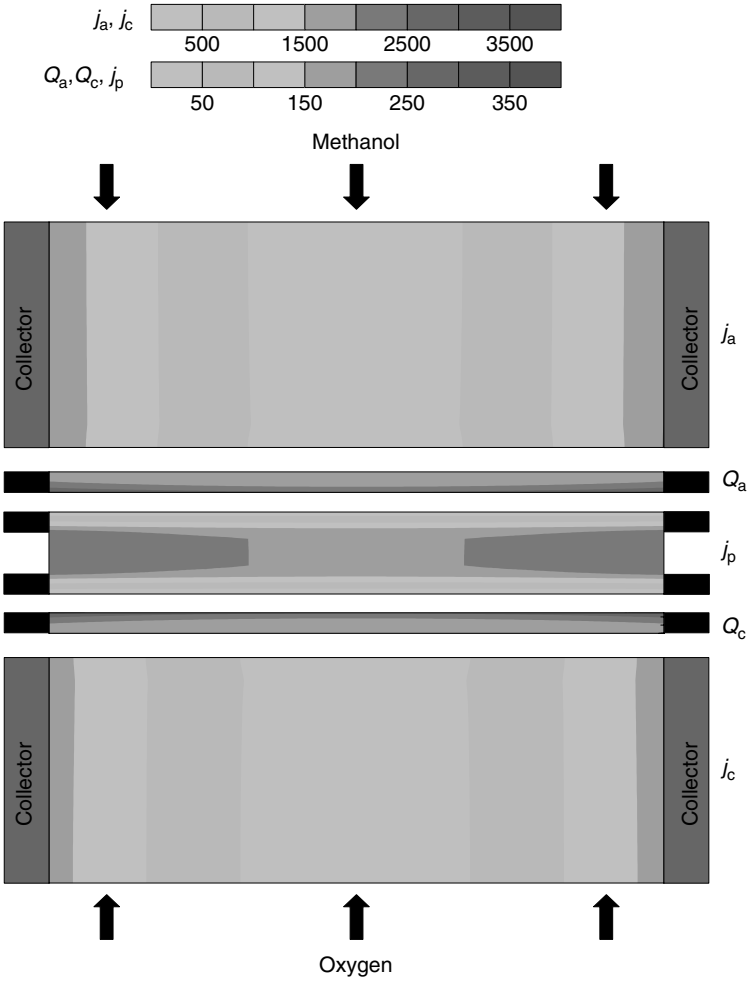


Fig. 23 Cell element with embedded current collectors. Units are given in caption to Fig. 22.

surprising that a similar idea of meshed electrodes gave excellent results [193]. The idea of embedded collectors is equally applicable to both electrodes of hydrogen PEFC.

8.2.4.5.2 **Gas-fed DMFC with Long Feed Channel** The next example is a gas-fed DMFC with a single meander-like channel on both sides. Cell operation conditions and parameters are specified in Table 5.

The maps of methanol and oxygen concentrations, electrochemical reaction rates, membrane phase potential and proton current density are shown in Fig. 24. The mean current density is 0.3 A cm^{-2}. Several interesting features are seen.

In the backing layer, methanol is transported mainly along x, and methanol-depleted (shaded) regions form in front

Tab. 5 Cell operation conditions and parameters for simulation of gas-feed DMFC

	Anode side	*Cathode side*
Cell temperature T [°C]	110	110
Inlet gas pressure [atm]	1.5	2
Inlet gas velocity [cm s^{-1}]	100	100
Oxygen molar fraction at the inlet	–	0.2
Water vapor molar fraction at the inlet	0.79	0.01
Nitrogen molar fraction at the inlet	–	0.79
Methanol molar fraction at the inlet	0.2	–
CO_2 molar fraction at the inlet	0.01	–
Exchange current density i_* [A cm^{-3}]	1×10^{-2}	1×10^{-5}
Transfer coefficient α	0.5	2.0
Reaction order γ	1.0	0.5
Membrane phase conductivity [S cm^{-1}]	0.0315	0.0315
Carbon phase conductivity [S cm^{-1}]	40	40
Catalyst layer thickness [µm]	10	10
Backing layer thickness [µm]	100	100
Membrane thickness [µm]	20	20
Feed channel length [cm]	80	80
Feed channel height [cm]	0.1	0.1

of the current collectors. The reaction rate, cf. Eq. (130), depends linearly on the methanol concentration ($\gamma = 1$, Table 5), which leads to a lower reaction rate in these regions.

Since the conductivity of the carbon phase is large, the electrode potentials keep track of the variation of membrane phase potential φ_{el}. Note that φ_{el} varies slowly along the catalyst layer/membrane interface (cf. Fig. 24).

Figure 25 shows the profiles of oxygen concentration along the channel for the current densities 0.1, 0.2, and 0.3 A cm^{-2}. At 0.3 A cm^{-2} oxygen concentration in the last (9th) element is reduced by a factor of two compared to the value at the inlet. To understand the law of feed gas consumption, a simple theory was developed [194], as described in the next section.

8.2.4.5.3 Theory of Feed Gas Consumption

Consider the cathode channel (similar arguments are applicable to the anode channel of DMFC or hydrogen PEFC). For simplicity we assume that (1) the catalyst layer is thin enough, so that there are no voltage losses associated with proton transport across the layer; and (2) the diffusion losses of oxygen in the backing and catalyst layers are negligible.

Under these assumptions the electrode potential and the rate of the electrochemical reaction are constant across the catalyst layer and the local current density is

$$j(z) = lQ_c(z) \approx li_c \frac{c_{O2}}{c_{O2ref}} \exp\left(\frac{\alpha_c F}{RT}\eta_c\right) \tag{143}$$

(cf. Eq. 129), where we assumed a reaction order $\gamma_c = 1$. (Note that i^*, used in Sect. 8.2.3 is $i^* = li_c c_{O2}/c_{O2ref}$.)

The oxygen concentration in the channel is defined by the continuity equation

$$v_c^0 \frac{\partial c_{O2}}{\partial z} = -\frac{S}{nF}\frac{j(z)}{h_c} \tag{144}$$

Here we assumed that velocity does not change dramatically with z, so that the estimate $v_c(z) \approx v_c^0$ is valid. The latter assumption is justified if v_c^0, the velocity at the inlet, is large enough. In addition we will assume, hereafter, that the cathode overvoltage η_c only weakly depends on z and that the only z-dependent quantity in Eq. (143) is c_{O_2}. As simulations show, such assumption works amazingly well but only if the membrane is well humidified.

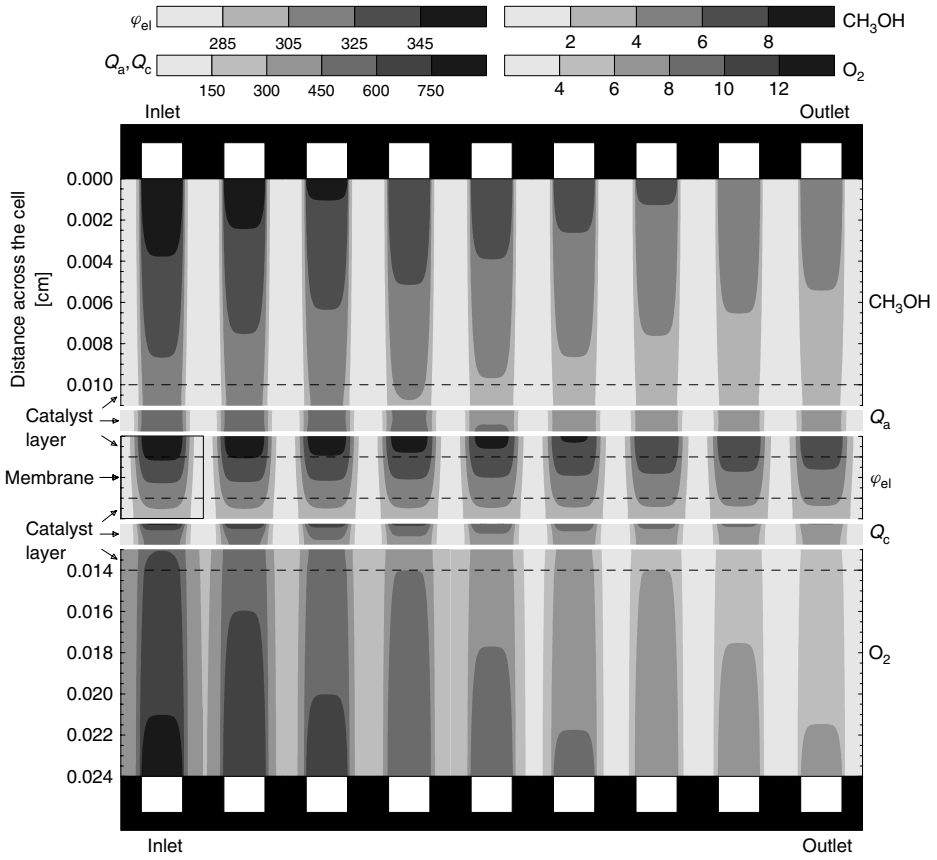


Fig. 24 Map of the gas feed DMFC. Shown are methanol and oxygen molar concentrations (10^{-6} mol cm^{-3}), the rates of electrochemical reactions (A cm^{-3}) and membrane phase potential (*mV*). Since carbon phase potential is practically constant, φ_{el} reproduces the shape of electrode potentials.

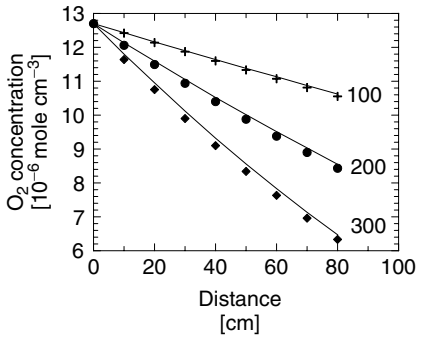


Fig. 25 Oxygen concentration along the channel (10^{-6} mol cm^{-3}) for indicated values of the mean current density in the cell (mA cm^{-2})

What does this assumption mean? It essentially implies that the membrane phase is equipotential along the channel. Of course, generally speaking, it is not! But the ohmic losses in the membrane in a fully humidified case are small compared to the reaction losses, so that the local violations of the equipotentiality will be hardly noticeable. This simplifying assumption allows getting a fully analytical solution to the problem.

Equations (143) and (144) allow us to eliminate $j(z)$. Solving the resulting equation for c_{O2} with the boundary condition $c_{O2}(0) = c^0_{O2}$ one finds

$$c_{O2}(z) = c^0_{O2} \exp\left(-\frac{z}{L_1}\right) \qquad (145)$$

where

$$L_1 = \frac{nF}{S}\frac{h_c v^0_c c_{O2\mathrm{ref}}}{l_c i_c} \exp\left(-\frac{\alpha_c F}{RT}\eta_c\right) \qquad (146)$$

is a characteristic length of oxygen consumption for $\gamma_c = 1$ (the solution for $\gamma_c \neq 1$ is given in [194]). Interestingly, for $\gamma_c < 1$, L_γ is a length of complete consumption of the feed, that is, $c_{O2}(L_\gamma) = 0$. In Fig. 25 the simulated and theoretical profiles of oxygen concentration along the channel (for conditions listed in Table 5) are compared.

Integrating Eq. (143) with Eq. (145) over z one finds the voltage loss due to oxygen exhaustion [195]:

$$\Delta V_c = \frac{RT}{\alpha_c F} \ln\left(-K_1 \ln\left(1 - \frac{\overline{j}}{j_{\lim}}\right)\right),$$

$$K_1 = \frac{nF}{S}\frac{h_c v^0_c c_{O2\mathrm{ref}}}{l_c i_c L} \qquad (147)$$

where $\overline{j} = (1/L)\int_0^L j(z)\mathrm{d}z$ is the mean current density in the cell and

$$j_{\lim} = \frac{nF}{S}\frac{h_c v^0_c c^0_{O2}}{L} \qquad (148)$$

is a limiting current density due to oxygen consumption in the channel. For a general expression at arbitrary γ_c see Ref. 195.

Figure 26 shows the voltage current curves of the cell for three indicated values of channel length L (cm). The shorter the channel, the better is the cell performance. Note that in the case of a short channel the voltage loss is lower in the whole range of current densities, not only near the limiting value $j_{\lim}$. Experimental verification of the influence of oxygen exhaustion on a voltage current curve of the cell [195] agreed well with the predictions of the theory.

In practice, the cell is often operated under constant stoichiometric flow ratio

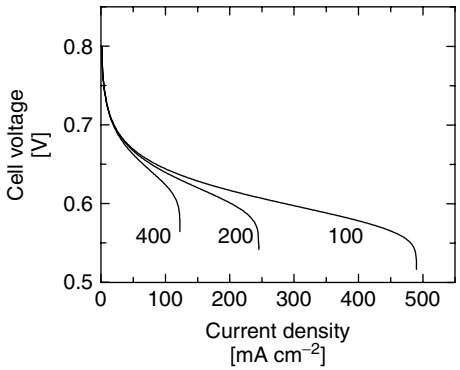


Fig. 26 Voltage–current curves of the cell with a conventional meander channel for the three indicated values of the channel length (cm).

and constant pressure of feed gases. This means that the inlet velocity (and hence j_{lim}) increases proportionally to mean current density. Combining the formulas above it is easy to obtain

$$L_1 = -\frac{L}{\ln\left(1-\frac{1}{\lambda}\right)} \tag{149}$$

which shows that L_1 does not change if λ is constant. Therefore, the profile of feed gas concentration does not vary with the current in that case. Further analysis leads to remarkably simple formulas for the profiles of feed (oxygen) concentration and local current density along the channel [196]

$$c_{O2}(z) = c^0_{O2}\left(1-\frac{1}{\lambda}\right)^{z/L} \tag{150}$$

$$\frac{j(z)}{\bar{j}} = \left(-\lambda\ln\left(1-\frac{1}{\lambda}\right)\right)\left(1-\frac{1}{\lambda}\right)^{z/L} \tag{151}$$

Recently, the profiles of $j(z)$ have been directly measured [197]. The results [197] obtained in conditions of a well humidified membrane are in excellent agreement with this simple theory [199].

8.2.4.5.4 Polarization Curves of Individual Segments of a Cathode Channel: S-shape Behavior as a Signature of Oxygen Starvation

The experiment with segmented flow field clearly demonstrates that the segments located far from the channel inlet exhibit a striking feature: their individual I–V curves are S shaped. Simple extension of the model above enables us to explain the effect.

Taking into account voltage loss due to oxygen transport through the backing layer, one arrives at the following expression for polarization voltage of the cathode side [198]

$$\frac{\eta}{b} = \ln\left(\frac{j}{j_*}\right) - \ln k - \ln\left(1-\frac{j}{j_D}\right) \tag{152}$$

where $j_* = 2\sigma_{\mathrm{el}}b/l$ is characteristic current density, $k = (li_*/j_*)(c_{O2}/c_{O2\mathrm{ref}})$ and

$$j_D = \frac{nFD_bc_{O2}}{Sl_b} \tag{153}$$

is the limiting current density due to oxygen transport through the backing layer. Here D_b is oxygen diffusion coefficient in the backing layer of a thickness l_b. Relation (152) describes the limiting behavior of cell polarization curve: as j approaches j_D, η tends to infinity.

Equation (152) contains oxygen concentration in the channel $c_{O2} = c_{O2}(z)$, which obeys Eq. (144). Solution to Eqs. (152) and (144) at the assumption z independent η is given by [199]

$$\frac{j}{j^0_D} = a\exp\left(-\frac{a\tilde{z}}{\beta}\right) \tag{154}$$

where $\tilde{z} = z/L$ (along-the-channel coordinate given in the units of the portion of the total length of the channel) a is a function of polarization voltage:

$$a = \frac{1}{1+\frac{1}{K}e^{-\eta/b}} \tag{155}$$

and

$$j^0_D = \frac{nFD_bc^0_{O2}}{Sl_b}, \quad K = \frac{Sl_bli_*}{nFD_bc_{O2\mathrm{ref}}},$$

$$\beta = \frac{l_{\mathrm{b}}hv^0_{\mathrm{c}}}{D_bL} \tag{156}$$

are constants.

Equation (154) gives the local polarization curve of individual segment at a distance z from the channel inlet. The set of curves for different $\tilde{z}$ is shown in Fig. 27.

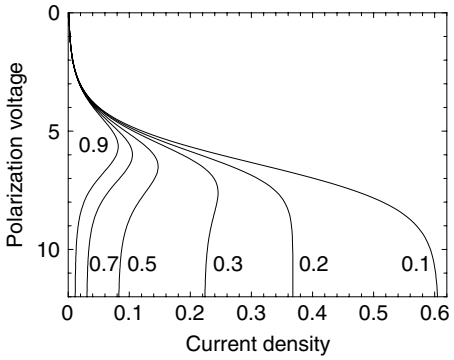


Fig. 27 Local polarization curves at the indicated distances from the inlet of the oxygen channel (0 is at the inlet, 1 is at the outlet). Shown is the local polarization voltage versus local current density. All values are dimensionless.

We see that this simple model well describes the characteristic S-shaped profiles of local polarization curves far from the channel inlet, detected in experiments [197, 200]. Physically, these maxima result from the effect of ''oxygen starvation''. Qualitatively, Eq. (154) shows that local current is a product of two factors. The preexponential factor a describes the growth of local current with the increase in η due to exponential dependence of the rate constant of ORR on η (Tafel law). The second (exponential) term in Eq. (154) describes oxygen consumption upstream from the given point $\widetilde{z}$.

Experiments [197, 200] are performed under constant inlet velocity, that is, at a fixed β. Typically, $\beta \simeq 0.1-1$, whereas parameter K is very small, $K \simeq 10^{-4} - 10^{-5}$. If η is small we have $K \ll \exp(-\eta/b) \simeq 1$, $a \simeq K \ll 1$ and the exponent in (154) is close to 1 for all $\widetilde{z}$. The dependence $j(\eta)$ is then determined by the first factor: $j/j_D^0 \simeq a$. Using (155) and solving for η we find $\eta/b = -\ln\left(K\left(j_D^0/j - 1\right)\right)$, which grows monotonously with j.

However, with the growth of η the exponent in (154) turns on. At certain z the exponent decreases faster, than increases the preexponential factor a. Physically, the amount of oxygen available at $\widetilde{z}$ decreases faster, than increases the rate constant of ORR and local current starts to drop. It easy to find the ''turning point'', η_*, for each local polarization curve from equation $\partial j/\partial\eta = 0$. Differentiating (154) and solving the resulting equation for η we find

$$\frac{\eta_*}{b} = -\ln\left(K\left(\frac{\widetilde{z}}{\beta} - 1\right)\right) \qquad (157)$$

The respective local current density is $j_*/j_D^0 = \widetilde{z}\exp(-1)/\beta$. Clearly, η_* exists if $\widetilde{z} > \beta$. The point $\widetilde{z} = \beta$ thus separates the domains of monotonic and nonmonotonic local polarization curves [201]. Note that if for some reason the local current in the ''monotonic'' domain is limited, the cell voltage exhibits oscillations [201].

8.2.4.5.5 Differential Pressure Flow Field

As it was shown already, an interdigitated flow field, which utilizes pressure gradient forces between the inlet and outlet channels, has distinct advantages

Tab. 6 Conditions and parameters for DMFC with differential pressure flow field

	Anode side	*Cathode side*
Inlet gas pressure [atm]	1.5	2 (1.95 red)
Membrane phase conductivity [S cm^{-1}]	0.034	0.034
Feed channel length [cm]	19.35 (red)	19.35 (red)

over the conventional flow field [167, 168, 202–206].

However, a pressure gradient can be applied in a conventional flow field as well [177]. We present here an example, which shows how one can facilitate water removal from the cathode, using this principle. Consider gas-fed DMFC with the three parallel meander-like channels on both sides of the cell (cf. Fig. 28). On the anode side the pressure in all channels is the same, whereas on the cathode side the pressure in the Red channel is lower than in the other two.

Most of the basic conditions and parameters of the simulation are the same as in Table 5; those which are different are given in Table 6. The map of the cell is presented in Fig. 29. Application of a pressure gradient changes the distribution of water vapor concentration at the cathode dramatically. Due to the pressure gradient water accumulates in front of the Red channel. This leads to a significant flux of water into this channel (cf. Fig. 30). Under current density 0.4 A cm^{-2} the flux of water is about 0.1 A cm^{-2} (cf. Fig. 30). The Red channel, therefore, collects water.

In a multichannel flow-field design the channels with ''normal'' and lowered pressure may alternate and one may expect that the latter will effectively remove excessive water from the cathode compartment. Such a flow field may have advantages over the interdigitated flow field since gas flow in the water-draining (Red) channels is externally driven and can be varied to reach optimal conditions.

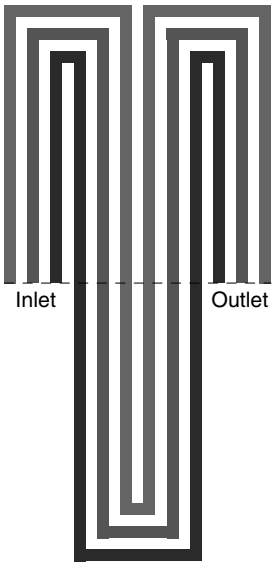


Fig. 28 A sketch of the flow field. On both sides of the cell the flow field is formed by three parallel meanders. On the anode side in all channels pressure is the same. On the cathode side pressure in the Green and Blue channels is the same; in the Red channel it is reduced. (Red channel is between Green and Blue).

8.2.4.5.6 **Hydrogen Cell (PEFC)** To apply the Q3D model to hydrogen PEFC, one has to modify the gas composition, the diffusion coefficients and the reaction rate parameters on the anode side. The parameters used for the PEFC simulation are listed in Table 7 (those parameters which are not shown here are the same as in Table 5). Since hydrogen oxidation

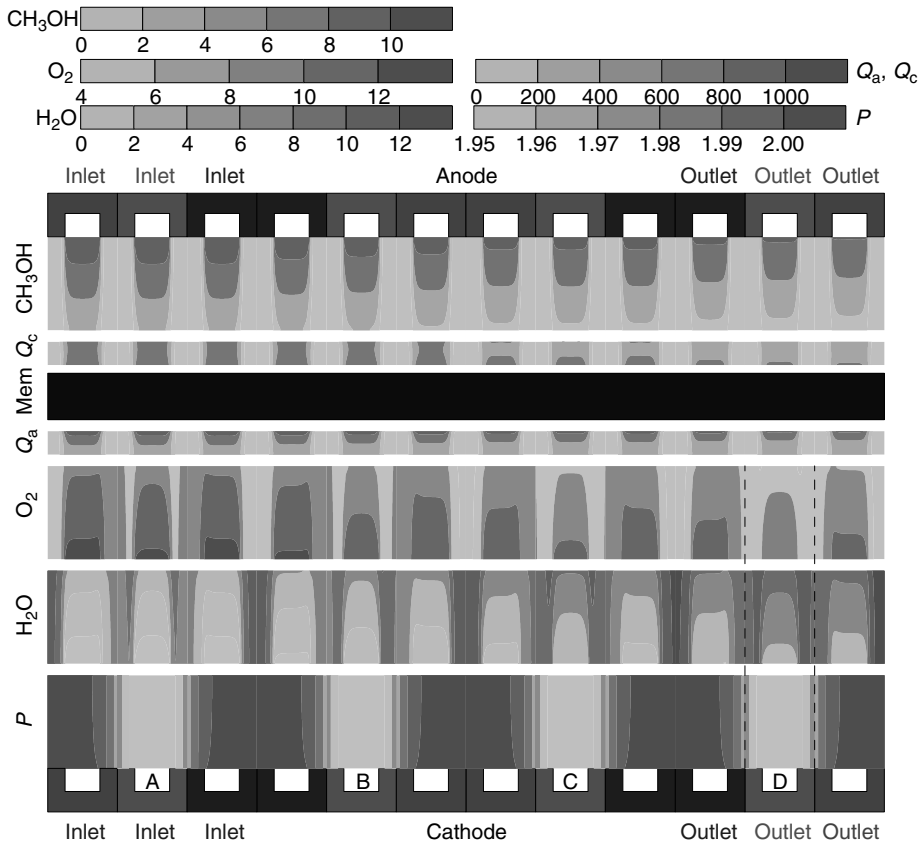


Fig. 29 Map of the gas-feed DMFC with differential pressure flow field. Shown are methanol, oxygen, and water vapor molar concentrations (10^{-6} mol cm^{-3}), the rates of electrochemical reactions Q_a, Q_c (A cm^{-3}), and the pressure in the cathode compartment (atm).

is much faster than oxidation of methanol the exchange current density at the anode of PEFC is several orders of magnitude larger than in DMFC.

The sketch of the simulated PEFC is shown in Fig. 31. The maps of parameters in the leftmost and rightmost cell elements are displayed in Fig. 32.

Due to small mass, hydrogen diffusivity is much higher, than that of methanol and oxygen. Near the inlet H_2 almost uniformly fills the backing and catalyst layers (cf. Fig. 32). Near the outlet, however, hydrogen-depleted regions form in front of the current collectors. On the cathode side the distribution of reaction rate reflects the distribution of oxygen concentration and η_c.

On the anode side, the reaction rate follows the distribution of η_a. Figure 32 shows that the reaction rate is high near the membrane/catalyst layer interface, where η_a is high. The difference between the profiles of reaction rates across the anode and the CCL is tremendous.

Tab. 7 Conditions and parameters for hydrogen–oxygen fuel cell

	Anode side	*Cathode side*
Cell temperature T [°C]	85	85
Inlet gas pressure [atm]	1.5	2
Inlet gas velocity [cm s^{-1}]	50	50
Oxygen molar fraction at the inlet	–	0.2
Water vapor molar fraction at the inlet	0.40	0.01
Nitrogen molar fraction at the inlet	–	0.79
Hydrogen molar fraction at the inlet	0.45	–
CO_2 molar fraction at the inlet	0.15	–
Exchange current density i_* [A cm^{-3}]	1.4×10^5	1×10^{-5}
Transfer coefficient α	0.5	2.0
Reaction order γ	1.0	1.0
Red channel length [cm]	25.8	25.8

Figure 33 shows the profiles of reaction rates across the catalyst layers in the inlet element (along the white line in Fig. 32). In this example, the reaction rate on the cathode side is almost constant along x, whereas on the anode side it rapidly grows with x. Most of the anode catalyst layer thickness is not used for reaction, except for the thin sublayer near the membrane, with the thickness of the order of 2–3 μm.

Physically, the reason for the dramatic difference between performances of cathode and anode active layers is the exchange current density i_a: at the anode the latter is 10 orders of magnitude higher than at the cathode [6]. Due to the large i_a, the electrode potential η_a is small. The anode of PEFC, hence, operates in the "linear" regime, when both exponential terms in the Butler-Volmer equation can be expanded [178]. This leads to exponential variation of η_a across the catalyst layer with the characteristic length (in the exponent)

$$l_a^* = \sqrt{\frac{\sigma_{el} b^a}{2 i_*^a} \left(\frac{c_{H_2 ref}}{c_{H_2}} \right)^{\gamma}}, \qquad (158)$$

which is similar to the well-known reaction-penetration depth in the context of GEDs [11, 106]. Figure 34(a) shows the exponential fit of the simulation data. The slope of the fitting curves is in excellent agreement with that given by Eq. (158). Thus the concept of reaction-penetration depth is valid for PEFC anodes with long feed channels at practically interesting current densities.

8.2.5 From Understanding to New Designs

We have shown a number of findings that can be taken into account to

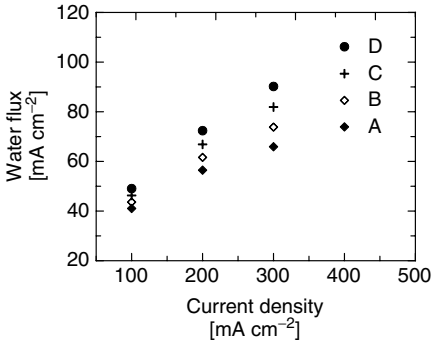


Fig. 30 Water flux (mA cm^{-2}) to the Red channel on the cathode side (molar flux of water is multiplied by 2F). The cross sections A, B, C, and D are shown in Fig. 29.

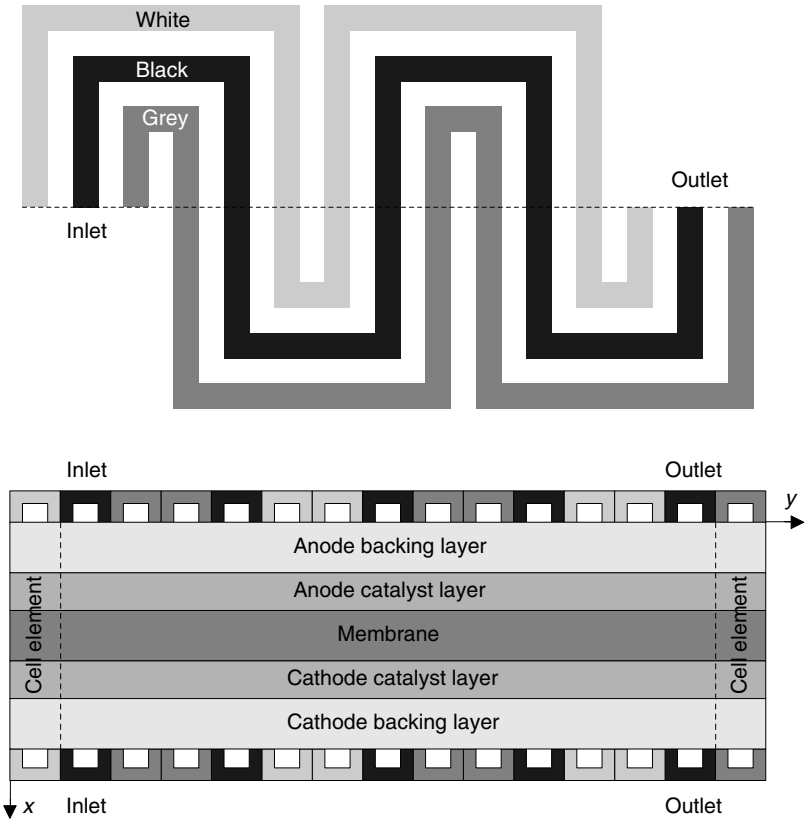


Fig. 31 Sketch of the simulated fuel cell. On both sides of the cell the flow field is formed by three parallel meander channels (shown on top). The cell cross section is shown below. The computational domain consists of 15 identical elements (one element is depicted by dashed lines).

different extents, after proper experimental verification, in an optimized design of the cell or cell components. Let us list them in the sequence of a rising potential for applications today.

In the first part we have drawn some guidelines in understanding the physics of membrane performance. These were mechanisms of the increase of the conductance of a single water-filled pore (at decreasing activation energy) with the increase of the pore size. We have seen how the single-pore conductance translates into the performance of the continuous proton-conducting network. A relationship was shown between the psd and the ability of a membrane to return water hydraulically to the region adjacent to the anode, which electroosmosis tends to dehydrate. Admittedly, this knowledge does not give us immediate suggestions on how to fabricate better membranes. It can stimulate our imagination, however. Imagine, for example, a membrane with two separate hydrophilic pore networks, one lined with acid groups (proton-conducting pores) and one without acid groups, but big channels for

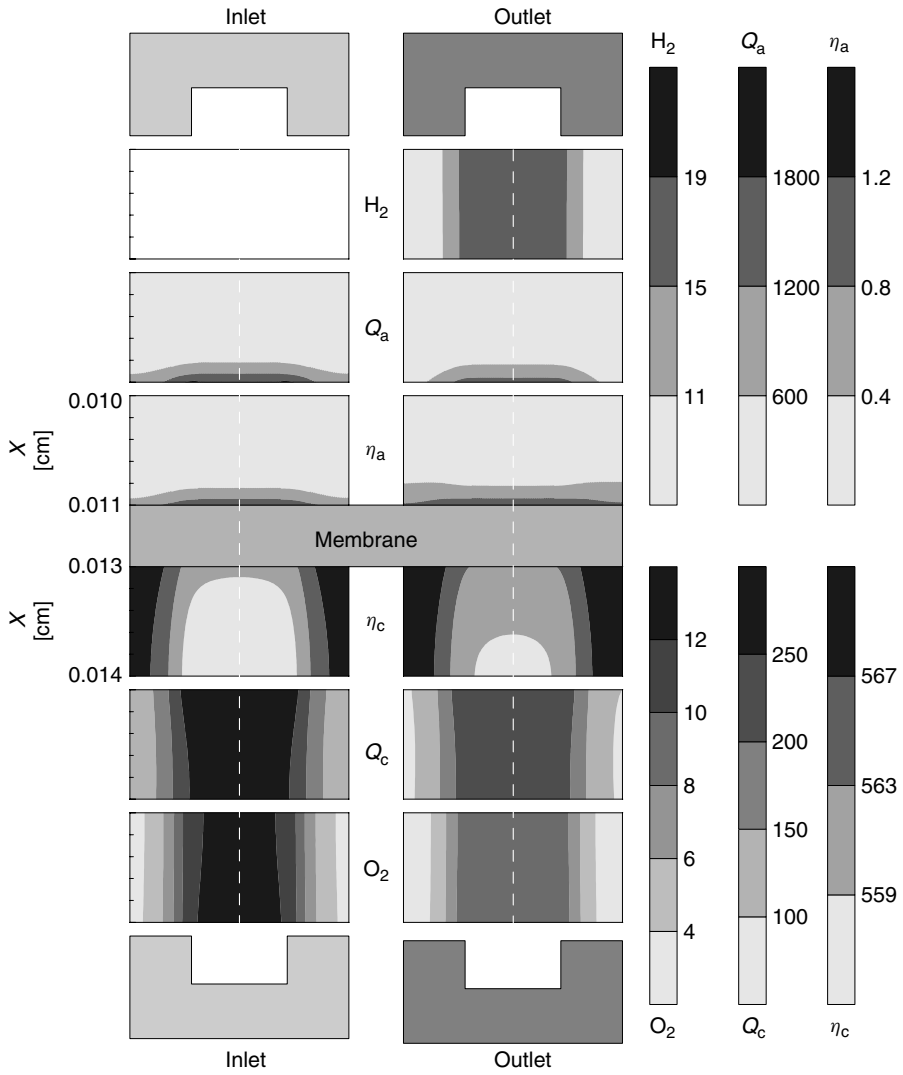


Fig. 32 Maps of hydrogen concentration (10^{-6} mol cm^{-3}), reaction rates (A cm^{-3}), electrode potentials (*mV*), and oxygen concentration (10^{-6} mol cm^{-3}) in the anode and cathode catalyst layers. Shown are the maps for the elements at the inlet of the White channel and at the outlet of the Grey channel (Fig. 31). Mean current density in the cell is 200 mA cm^{-2}.

hydraulic water flux – the structure would allow control of electroosmosis and backflow independently. If realized, theory can help to optimize parameters of such a membrane. Still, staying away from science fiction, one can systematically check the effect of the flexibility and length of the side chains, as observed in the simulations. A more precise morphological picture of the membrane should be verified by further experiments (beyond the established correlations between the

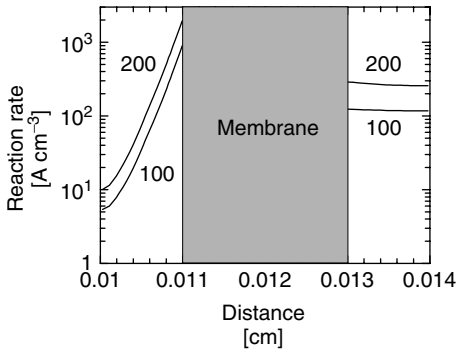


Fig. 33 The profiles of reaction rates across the catalyst layers in the first element (at the inlet of White channel). Plotted are the profiles along the dashed line, shown in Fig. 32.

macroscopic and microscopic swelling). Focusing on the bottlenecks for the proton transfer, it is suggested that the key measure to facilitate it would be making aqueous nanoscale necks between the unit cells of the quasicrystalline structure *slightly larger*, without ruining the elastic matrix of the ionomer. But a real design-oriented approach toward membrane improvement would require a concerted undertaking which incorporates the knowledge of chemical details on the polymer–water (or other solvents) interaction [207–209], the insight gained by the structural analysis of the ionomer, and MDs simulations; all this should be combined with what we understand about the membrane operation in situ. Furthermore, one cannot speculate about optimized membrane properties, without considering elastic properties of the membranes on the level of microstresses and resistance to damage, polymer statistical properties and its ability to denature. The concepts provided above help to understand the membrane operation in a fuel cell; they tell us how the membrane should be operated or which type of membrane should be used in order to extract the best possible performance from the cell. A couple of straightforward suggestions concerning the observation and the study of the currents limited by dehydration were presented in the summary section of the membrane part together with several other hints.

The practical outcome of the second part is the dimensionless performance diagram of the CCL, which may be extended with minimal modifications onto the anode. This diagram allows fixing an optimum thickness of the layer subject to the relevant current regime. To use this diagram for each particular layer, we need to know the proton conductivity of the layer, its gas permeability, and parameters of the equation for reaction rate in the layer. This requires a lot of systematic experimental work. Parameterization of the catalyst layer performance as a function of the polymer-electrolyte content was based on certain expressions coming from the lattice percolation theory. It gives only a hint of what these dependencies *could* look like and whether such simple ideas could, in principle, rationalize the measured families of the current–voltage plots. Although they appear to be able to reproduce existing experimental data, the laws or at least parameters of the lattice percolation theory might be essentially different in a polydisperse complex metal electrolyte mixture [104].

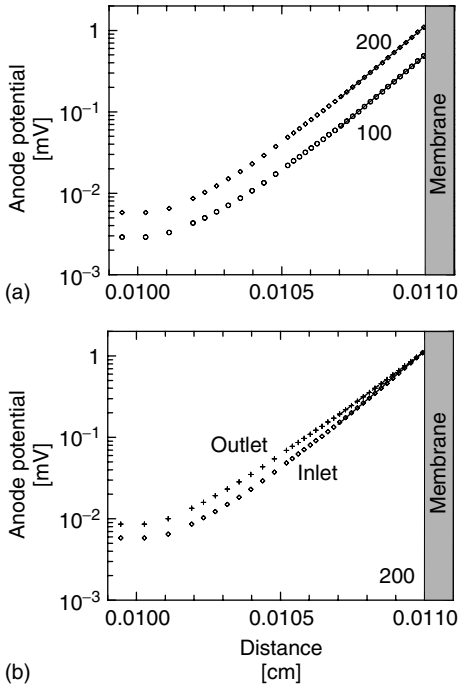


Fig. 34 (a) The profile of electrode potential η_a across the anode catalyst layer at the inlet of white channel for indicated values of mean current density in the cell (mA cm^{-2}). η is shown along the same, as in Fig. 33 line. (b) The profiles of electrode potentials across the anode catalyst layer in the "first" and "last" cell elements (at the inlet of White channel and at the outlet of Grey channel, respectively). Mean current density: 200 mA cm^{-2}.

Despite all the complications, catalyst layer theory provides a few clear instructions for improvements. Composition and thickness can be readily adjusted, following recent optimization studies. The theory was used, for instance, to study upgraded design with nonuniform compositions [123, 124]. Performance enhancements measured for correspondingly fabricated functionally graded CCLs are in line with predictions of the theory.

The theory-based estimate, that Pt utilization in conventional three-phase composite CCLs reaches only 10–20% of its full potential, gives rise to optimistic speculations on how much better catalyst layers could be made. Exploring these reserves requires innovative fabrication procedures. Such efforts could be facilitated decisively, however, by upgrades in the theoretical understanding of catalyst layer functioning. Aspects of partial liquid water saturation, water formation/transformation and two-phase transport in porous structures should be included in the macro-homogeneous electrode theory. At microscopic or atomistic resolution, on the other hand, improved theoretical understanding of relevant reaction mechanisms in nanoparticle systems could reveal routes toward nanostructured electrocatalysts with highest activities.

The suggestions that come out from the third part are more straightforward:

1. In the cells of conventional flow-field geometry there are shaded zones in the catalyst layer, *behind* the current collector plates or *in front* of the gas channels [18], depending on the permeability and conductivity of the backing layers. This asks for saving precious catalyst material by removing it from the reaction-dead zones. However, the mere existence of dead zones immediately suggests a better design of embedded current collectors which provides a much more homogeneous distribution of events in the cell, in which the dead zones do not emerge.
2. The use of the parallel meanders with differential pressure should help water removal from the channels of the cathode.

3. The characteristic or even finite length of feed consumption of the fuel in the gas supplying channels and the existence of the corresponding limiting currents was predicted by the simulations and theory [194] and verified experimentally, cf. [195]. For a given regime of feeding and the channel structure, the theory helps to calculate the characteristic length of the feed consumption, which is instructive for the cell design, and, last but not least, helps to rationalize the existing empirical tricks which are used to avoid cell starvation. Theoretical understanding of the starvation laws is in line with the recent experimental findings of F. Buechi's group at the Paul Scherrer Institute, Villingen. They designed a cell having a gradual decrease of platinum loading along the channel from the inlet to the outlet. Using platinum loading patterns optimized for the target stoichiometry was shown to be a cost-efficient solution.

More examples could be drawn, if not limited by the length of this chapter, as the field is rapidly developing. To facilitate its faster success, a much higher level of cooperation between the theory, electrochemical experiments (also in the nanoscale range [187, 188, 210]), and modified stack testing is vital. Such cooperation, however, requires a different level of openness between academia and industry, than exists in the current gold rush atmosphere of PEFC/DMFC research and development.

Symbols – Membrane Water Management, Section 2

c_w	Concentration of water, $\approx$55 mol l^{-1}
$D(w_x)$	Water diffusion coefficient (dimensionless)
j_w, j_{wa}, j_{wc}	Water flow (membrane, anode, cathode), A cm^{-2}
j_p	Proton current density, A cm^{-2}
j_{pc}	Critical proton current density, A cm^{-2}
j_m	Characteristic parameter, $j_m \propto j_{pc}$, A cm^{-2}
$K(w_x)$	Hydraulic permeability, nm^2
L	Membrane thickness, μm
$n(w_x)$	Electroosmotic coefficient
P^g, P^l, P^c	Gas, liquid, capillary pressure, atm
r	Pore radius, nm
r_1	First moment of the pore radius distribution, nm
$\Re$	Membrane resistance, Ω cm
U	Membrane potential drop, V
w	Water content (volume fraction) in the membrane
w_c	Percolation water content
w_s	Saturation water content
x	Coordinate perpendicular to membrane surface, μm

Greek symbols

γ	Surface tension of water in membrane pores, J m^{-2}
μ	Viscosity, mJ s cm^{-2}
ξ	Tortuousity factor
$\sigma_{el}(w_z)$	Membrane conductivity, S cm^{-1}
φ_{el}	Local membrane potential, V
"psd"	Pore size distribution
"REV"	Representative elementary volume

Symbols – Electrode Performance, Section 3

b	Apparent Tafel slope, $b = RT/\alpha_c n_t F$, V
$\tilde{b}$	Tafel parameter in linear regime, $\tilde{b} = RT/n_t F$, V
$\bar{c}_{O_2}$	Oxygen concentration at diffusion layer interface, mol l^{-1}
C_{dl}	Double-layer capacitance, Fcm^{-2}
g	Ratio of O_2 diffusion versus proton conductivity, $g = I/\sigma b$
I	Characteristic parameter of O_2 diffusion, A cm^{-2}
i^*	Exchange current density parameter, A cm^{-2}
j_p	Local proton current density, A cm^{-2}
j_0	Proton current density at interface to membrane, A cm^{-2}
l	Catalyst layer thickness, µm
l_p	Reaction penetration depth (proton transport), µm
L_p	Renormalized reaction penetration depth, µm
P_{O_2}	Oxygen partial pressure, atm
p	Nondimensional pressures, $p = P_{O_2}/P^*_{O_2}$
$\Re_c$	Linear electrode polarization resistance, Ω cm^2
$\Re_{diff}$	Differential resistivity, Ω cm^2
S_0	Effective catalyst\|electrolyte contact area, cm^{-1}
x	Coordinate perpendicular to catalyst layer surface, µm
χ	Nondimensional coordinate, $\chi = x/l$
X_{el}, X_m, X_c	Volume portions (electrolyte, metal, percolation threshold)
$Z(\omega)$	Complex impedance, Ωcm^2

Greek symbols

Γ	Dimensionless parameter, $\Gamma = (i^* I/j_0^2)\exp(\eta/b)$
δ_{eff}	Thickness of active part of the catalyst layer, µm
η	Local electrode potential, V
η_0	Catalyst layer overpotential, V
σ_{el}	Specific proton conductivity, S cm^{-1}
σ	Proton conductance per surface area, $\sigma = \sigma_{el}/l$, S cm^{-2}
σ_p	Mixed conductance parameter, S cm^2 (def. in Eq. (52))
Σ_p	Renormalized conductance, S cm^2
φ_{el}	Local electrolyte potential, V
φ_c	Potential of carbon/catalyst phase, V
$\chi_{ec}(\chi_d)$	Parameter of residual reactivity (diffusivity)
ω	Frequency of sinusoidal signal, s^{-1}
ω_p	Characteristic frequency (renormalized: Ω_p), s^{-1}

Symbols – Q3D Model, Section 4

c	Molar concentration, mol cm^{-3}
h	Channel height, cm
i	Exchange (volumetric) current density, A cm^{-3}

j	Local current density, A cm^{-2}
$\overline{j}$	Mean current density in a cell, A cm^{-2}
j_{lim}	Limiting current density due to feed gas exhaustion, A cm^{-2}
k^{p}	Hydraulic permeability, cm^2
K_1	Dimensionless constant in (147)
l	Catalyst layer thickness, cm
l_{a}^*	Characteristic thickness (158) (anode catalyst layer of PEFC), cm
L	Channel length, cm
L_1	Characteristic length of feed gas consumption, cm
M	Molecular weight, g mol^{-1}
M	Mass of the molecule, g
n	Number of electrons, participating in the reaction
$\mathbf{N}_i$	Molar flux of ith component, mol cm^{-2}s^{-1}
p	Pressure, g cm^{-1}s^{-2}
Q	Rate of charged particles production in electrochemical reaction, cm^{-3}s^{-1}
S	Stoichiometric coefficient
v	Celocity of the flow in the channel, cm s^{-1}
x, y, z	Coordinates, cm

Subscripts and superscripts

0	At the inlet
a,c	Anode, cathode
el	Membrane phase
M	Methanol
n	Carbon phase
O_2	Oxygen
ref	Reference value
w	Water vapor

Greek symbols

α	Butler–Volmer transfer coefficient
ς	Correction factor for porosity (binary diffusion coefficient).
γ	Reaction order
Λ	Kinetic coefficient
μ	Gas viscosity, g cm^{-1}s^{-1}
σ	Conductivity, S cm^{-1}
ρ	Gas density, g cm^{-3}
ξ_i	Molar fraction of ith component.
ψ	Correction factor for tortuosity (Knudsen diffusion coefficient).

For reasons not depending on the authors, this chapter has been with the publisher for more than four years. The core of the text was written in 2002; some amendments and updations were done in 2005.

Following the philosophy of this encyclopedia, we tried to collect and illustrate the ideas that form the basis for a physical modeling of cells and their components. This field, however, is young and its basis is being continuously extended. The recent almost threefold increase in oil prices and a growing awareness of the adverse environmental effects of the current energy use intensify the economic and political stress, which in turn induces dramatic growth of the interest in fuel cells. The data from www.scopus.com show that the number of publications on fuel cells grows exponentially. This exponential growth is a signature of emerging technology; it means rapid addition of new knowledge, new effects, and new important details to our understanding of how the cells function.

Below we briefly review the works, which highlight the most recent achievements in the field. Fuel cell modeling is a multidisciplinary problem, which extends its limits capturing new areas in the adjoining sciences. The review below is definitely not complete; the list of cited papers includes the works that came into the authors' view.

Membrane Models

Our section on membrane modeling focuses primarily on functional requirements and performance modeling of polymer electrolyte membranes (PEM) for fuel cells. There have not been decisive new insights for this type of ''device'' modeling since the writing of this chapter. The very basics of the pertinent understanding of the structure of Nafion and its water sorption properties are revisited in Ref. 211. The relevant phenomenology of various transport phenomena and reactions in proton conductors for fuel cell applications has been reviewed recently in Ref. 212. An encyclopedic survey of existing models of transport phenomena in membranes can be found in Ref. 213.

An avalanche-like rush of new research activities has evolved around new routes in membrane synthesis and strategies in theoretical and computational modeling that could facilitate a deliberate design of highly functionalized fuel cell membranes. Comprehensive reviews on membrane synthesis highlight principles and fabrication of new membranes particularly those applicable for DMFC [214], those that are feasible for operation at elevated temperatures (up to 200°C under ambient pressures) [215], and those that are based on various modifications of Nafion-type membranes [216].

Recent interest in the theory and simulation of PEMs has shifted toward understanding molecular mechanisms of proton transport, exploring the formation of phase-segregated membrane morphologies upon hydration, and developing multiscale modeling approaches with predictive capabilities. Despite rapidly growing computational resources and tools for studying complex many-particle systems, the full multiscale modeling of structure formation and transport in membranes, starting from first principles and leading all the way through to properties and performance of a membrane as a macroscopic device, is still a distant perspective. Insightful activities focus on structure formation in systems consisting of charged polymer molecules and water, utilizing full-atomistic molecular dynamics [217–219], lattice Monte Carlo [220], and coarse-grained mesodynamics based on self-consistent field theory [221, 222]. These studies explore the emergence of rodlike fibrils and of water-containing pathways upon hydration.

Calculations utilizing density functional theory, ab initio molecular dynamics, and equivalent valence bond approaches focus on unraveling structures and transport mechanisms in narrow pores and in the interfacial region between polymer and water. The purpose is to obtain a detailed molecular-level understanding of the coupling between the dynamics of charged, flexible side chains and the transport of mobile protons in the adjacent strongly bound water layers [223–225]. The question of whether high rates of proton transport could be attained in a minimal hydrated membrane that only retains strongly bound water molecules at the interface has been addressed recently in experiment [226] and ab initio quantum mechanical calculations [225]. In this context, it is also expedient to

refer to a number of studies that focus on the molecular-level understanding of mechanisms of lateral proton transport along biological membranes and lipid monolayer films [227–232]. Similarities with proton transport phenomena in bioenergetics could stimulate the design of advanced, bioinspired PEMs.

Models of Catalyst Layers

This chapter focuses mainly on macrohomogeneous approaches that could rationalize the effects of catalyst layer thickness and composition on catalyst utilization and performance. It is relevant for the operation of random three-phase composite layers, which are still the state of the art in fuel cell technology. It is widely recognized that activity and utilization of the catalyst will have to be understood at finer structural resolution in order to be able to reduce catalyst loadings to the levels that are acceptable for industrial applications. Determining the true extent to which catalyst is utilized in current layers remains a challenge for experimental diagnostics. Theoretical estimates suggest that the effectiveness of catalyst utilization could be improved more than 10-fold.

A number of fundamental studies explore catalyst activity at an atomistic scale. DFT calculations can reveal how the rates of surface processes depend on the local electronic structure of surface atoms [233–238]. Monte Carlo simulations and mean-field approaches can incorporate this information in order to rationalize the effects of nanoparticle sizes and surface structure on the overall rates of current conversion [233, 239]. Thereby the nontrivial dependence of reactivity on particle size could be explained.

Experiments suggest that agglomerates of carbon particles form distinct structural units in the fabrication process. Electrostatic effects control the spatial distributions of electrical potential and reaction rates in agglomerates. Corresponding effectiveness factors of agglomerates have been studied in Ref. 240 as a function of agglomerate radius, composition, and the charge-transfer coefficient.

Another vital step in rationalizing catalyst layer operation and long-term stability would be the understanding of structure formation during the fabrication process. Approaches similar to those employed in studying the structure formation in membranes could be utilized for this task. This work is currently in progress. Understanding how different conditions during fabrication and operation precipitate degradation of the catalyst layer structure would be helpful.

The role of the porous structure and partial liquid-water saturation in the catalyst layer in performance and fuel cell water balance has been studied in Ref. 241. As demonstrated, the cathode catalyst layer fulfills key functions in vaporizing liquid water and in directing liquid-water fluxes in the cell toward the membrane and cathode outlet. At relevant current densities, the accumulation of water in the cathode catalyst layer could lead to the failure of the complete cell. The porous structure controls these functions.

Diagnostics of the resulting structures and the corresponding effective properties of the catalytic layer continue to be the key issues. One could distinguish here the physical measurements of conductances and diffusivities of different phases of the composite, a difficult but not impossible task, which could indicate how good the percolation in each component is (see e.g.,

Refs. 242, 243). Another issue is monitoring the structure and its changes with degradation. We are not aware of any substantial results on this front. The relevant structures and their changes during operation are not only poorly predictable but they are also not sufficiently well characterized.

At high anodic overpotentials, methanol oxidation reaction exhibits strongly non-Tafel behavior owing to finite and potential-independent rate of methanol adsorption on catalyst surface [244]. The equations of Section 8.2.3 can be modified to take into account the non-Tafel kinetics of methanol oxidation. The results reveal an interesting regime of the anode catalyst layer operation featuring a variable thickness of the current-generating domain [245]. The experimental verification of this effect, however, has not yet been performed.

Models of Cells and Stacks

The approach described in Sections 8.2.3 and 8.2.4.5.3 was used to construct quasi-2D (Q2D) analytical and semi-analytical models of PEFC [246, 247] and DMFC [248, 249]. The Q2D model of a PEFC [246] takes into account water management effects, losses due to oxygen transport through the GDL, and the effect of oxygen stoichiometry. The model is fast and thus suitable for fitting; however, the systematic comparison of model predictions with experiment has yet not been performed. Q2D approaches have been employed to construct a model of PEFC performance degradation [250], to explain the instabilities of PEFC operation [251, 252] and to rationalize the effect of CO_2 bubbles in the anode channel on DMFC performance [253, 254].

The Q2D model of a DMFC [248] revealed a new effect: localization of current close to the inlet of the oxygen channel at small airflow rate. The effect is due to methanol crossover through the membrane: a system with crossover current is generated only close to the inlet of the oxygen channel, where oxygen is in excess (galvanic domain). In the rest of the cell, all available oxygen is consumed in a direct burning with permeated methanol. Further theoretical [255, 256] and experimental [257–260] studies of this regime have shown that the oxygen-depleted part of the cell works as an electrolytic cell: it consumes current produced in the galvanic domain to split methanol on the cathode side and to produce hydrogen on the anode side.

The Q2D approach was utilized to construct the models of PEFC [261–263] and DMFC [264, 265] stacks. Assembling the cells into a stack leads to another overhead: one needs to transport the current through the bipolar plates (BPs) separating individual cells. The respective voltage loss can be calculated on the basis of the model [262]. Generally, in-plane current in BPs can flow in any direction; the plates are thus true 2D objects [265]. A promising approach to stack modeling would be a combination of Q2D models for the description of individual cells with fully 2D models of current and heat transport in the BPs.

Progress in computer technology facilitates 3D modeling of cells [266, 267]. The 3D results help understand the role of 3D effects in cell function and to explore and optimize complex geometries of flow fields. Reduced models [268, 269], which utilize the idea of splitting the fully 3D or 2D problem into a number of problems of lower dimensionality seem to be very promising (this idea is similar to that described in Section 8.2.4). These models are

much faster, than their fully 3D counterparts, which is vitally important for cell optimization and massive comparison of model predictions with experiment.

Final Remarks

Fuel cells are probably the only viable alternative to traditional power sources based on direct combustion of fossil fuels (including uranium). The economy based on a direct combustion is as old as human society; it will inevitably come to an end with the exhaustion of fossil fuels. The idea of intelligent and very efficient two-step combustion with generation of charged particles as intermediates was suggested more than 150 years ago by Christian Friedrich Schoenbein (1838). It was immediately explored by Sir William Grove (1839), who was the first to demonstrate its great practical potential [270]. The fuels for this two-step combustion can, in principle, be produced from renewable sources like biomass, wind, and sunlight.

Fuel cells remain very attractive for both technical and aesthetical reasons. Probably nothing is more elegant and straightforward in the family of power generators. Thanks to nature, all pairs of half-cell reactions yield open-circuit voltage on the order of 1 V; learning how to use this potential in a most efficient way is a task that may change the history of civilization.

References

1. M. Eikerling, A. A. Kornyshev, A. A. Kulikovsky, *The Fuel Cell Rev.* **2004–2005**, 15.
2. M. W. Verbrugge, R. F. Hill, *J. Phys. Chem.* **1988**, *92*, 6778.
3. M. W. Verbrugge, R. F. Hill, *J. Electrochem. Soc.* **1990**, *137*, 886.
4. T. F. Fuller, J. Newman, in *Fuel Cells*, (Eds.: R. E. White, A. J. Appleby), Proc. Series PV 89-14, The Electrochemical Society, Pennington, 1989, p. 25.
5. D. M. Bernardi, M. W. Verbrugge, *AIC'hE* **1991**, *37*, 1151.
6. D. M. Bernardi, M. W. Verbrugge, *J. Electrochem. Soc.* **1992**, *139*, 2477.
7. T. E. Springer, T. A. Zawodzinski, S. Gottesfeld, *J. Electrochem. Soc.* **1991**, *138*, 2334.
8. T. E. Springer, S. Gottesfeld, in *Proceedings of the Symposium on Modeling of Batteries and Fuel Cells* (Eds.: R. E. White, M. W. Verbrugge, J. F. Stockel), PV 91-10, The Electrochemical Society, 1991, p. 197.
9. T. E. Springer, M. S. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1993**, *140*, 3513.
10. N. Nguyen, R. E. White, *J. Electrochem. Soc.* **1993**, *140*, 2178.
11. A. N. Frumkin, *Zh. Fiz. Khim.* **1949**, *23*, 1477.
12. O. S. Ksenzhek, V. V. Stender, *Dokl. Akad. NSSSR* **1956**, *107*, 280.
13. K. Mund, F. V. Sturm, *Electrochim. Acta* **1974**, *20*, 463.
14. I. Rousar, K. Micka, A. Kimla, *Electrochemical Engineering II, Part F*, Elsevier, Amsterdam, 1986.
15. R. P. Iczkowski, M. B. Cutlip, *J. Electrochem. Soc.* **1980**, *127*, 1433.
16. M. Eikerling, Yu. I. Kharkats, A. A. Kornyshev et al., *J. Electrochem. Soc.* **1998**, *145*, 2684.
17. M. Eikerling, A. A. Kornyshev, *J. Electroanal. Chem.* **1999**, *475*, 107.
18. A. A. Kulikovsky, J. Divisek, A. A. Kornyshev, *J. Electrochem. Soc.* **1999**, *146*, 3981.
19. A. A. Kulikovsky, J. Divisek, A. A. Kornyshev, *J. Electrochem. Soc.* **2000**, *147*, 953.
20. Ph. Colomban (Ed.), *Proton Conductors. Solids, Membranes and Gels – Materials and Devices*, Cambridge University Press, Cambridge, 1992.
21. T. Erdey-Grúz, *Transport Phenomena in Aqueous Solutions*, Adam Hilger, London, 1974.
22. N. Agmon, *Chem. Phys. Lett.* **1995**, *244*, 456.
23. O. Savadogo, *J. New Mater. Electrochem. Syst.* **1998**, *1*, 47.
24. M. Doyle, S. K. Choi, G. Proulx, *J. Electrochem. Soc.* **2000**, *147*, 34.

25. J. Ding, C. Chuy, S. Holdcroft, *Chem. Mater.* **2001**, *13*, 2231.
26. W. Grot, *Chem.-Ing.-Tech.* **1972**, *44*, 167.
27. G. A. Eisman, in *Diaphragms, Separators, and Ion-Exchange Membranes* (Eds.: J. W. Van Zee, R. E. White, K. Kinoshita et al.), Proc. Vol. 86-13, The Electrochemical Society, Pennington, 1986, p. 52.
28. A. Eisenberg, *Macromolecules* **1990**, *23*, 4098.
29. S. J. Paddison, *J. New Mater. Electrochem. Syst.* **2001**, *4*, 197.
30. M. Eikerling, S. J. Paddison, T. A. Zawodzinski Jr., *J. New Mater. Electrochem. Syst.* **2002**, *5*, 1.
31. J. J. Sumner, S. E. Creager, J. J. Ma et al., *J. Electrochem. Soc.* **1998**, *145*, 107.
32. H. K. Christensen, J. Fang, J. N. Israelachvili, *Phys. Rev. B* **1989**, *39*, 11750.
33. K.-D. Kreuer, S. J. Paddison, E. Spohr et al., *Chem. Rev.* **2004**, *104*, 4637.
34. S. Walbran, A. A. Kornyshev, *J. Chem. Phys.* **2001**, *114*, 10039.
35. E. Spohr, *Mol. Sim.* **2004**, *30*, 107.
36. P. Commer, C. Hartnig, D. Seeliger et al., *Mol. Sim.* **2004**, *30*, 755.
37. E. Spohr, P. Commer, A. A. Kornyshev, *J. Phys. Chem. B* **2002**, *106*, 10560.
38. P. Commer, A. G. Cherstvy, E. Spohr et al., *Fuel Cells* **2002**, *2*, 127.
39. M. Eikerling, A. A. Kornyshev, U. Stimming, *J. Phys. Chem. B* **1997**, *101*, 10807.
40. M. Eikerling, A. A. Kornyshev, A. M. Kuznetsov et al., *J. Phys. Chem. B* **2001**, *105*, 3646.
41. T. A. Zawodzinski Jr., C. Derouin, S. Radzinski et al., *J. Electrochem.Soc.* **1993**, *140*, 1041.
42. K. D. Kreuer, T. Dippel, W. Meyer, *J. Meier, Mater. Res. Soc. Symp. Proc.* **1993**, *293*, 273.
43. M. Eikerling, A. A. Kornyshev, *J. Electroanal. Chem.* **2001**, *502*, 1.
44. M. Eikerling, A. A. Kornyshev, *J. Electroanal. Chem.* **2002**, *528*, 196.
45. H. H. Voss, D. P. Wilkinson, P. G. Pickup et al., *Electrochim. Acta* **1995**, *40*, 321.
46. M. Cappadonia, J. W. Erning, S. M. Saberi Niaki et al., *Solid State Ionics* **1995**, *77*, 65.
47. Z. Luz, S. Meiboom, *J. Am. Chem. Soc.* **1964**, *86*, 4768.
48. L. I. Krishtalik, *Charge Transfer Reactions in Electrochemical and Chemical Processes*, Plenum, New York, 1986.
49. D. Marx, M. E. Tuckermann, M. Parinello, *J. Phys.: Condens. Matter* **2000**, *12*, A153.
50. U. W. Schmitt, G. A. Voth, *J. Chem. Phys.* **1999**, *111*, 9361.
51. A. A. Kornyshev, A. M. Kuznetsov, J. Ulstrup et al., *J. Phys. Chem. B* **1997**, *101*, 5917.
52. A. M. Kuznetsov, *Charge Transfer in Physics, Chemistry and Biology*, Gordon & Breach, Reading, 1995.
53. M. Eikerling, S. J. Paddison, L. W. Pratt et al., *Chem. Phys. Lett.* **2003**, *368*, 108.
54. T. D. Gierke, G. E. Munn, F. C. Wilson, *J. Polym. Sci. Polym. Phys. Ed.* **1981**, *19*, 1687.
55. G. Gebel, *Polymer* **2000**, *41*, 5829.
56. V. I. Gordelyi, A. Islamov, A. Kettner et al., to be published.
57. L. Rubatat, A. L. Rollet, G. Gebel et al., *Macromol.* **2002**, *35*, 4050.
58. O. Diat, S. Lyonnard, G. Gebel et al., *Physica B* **2004**, *350*, e959.
59. R. E. Baltus, *J. Membr. Sci.* **1997**, *123*, 165.
60. J. Divisek, M. Eikerling, V. Mazin et al., *J. Electrochem. Soc.* **1998**, *145*, 2677.
61. A. S. Ioselevich, A. A. Kornyshev, J. Steinke, *J. Phys. Chem. B* **2004**, *108*, 11953.
62. A. S. Ioselevich, A. A. Kornyshev, N. Marchal, to be published.
63. X. Ren, T. A. Zawodzinksi Jr., F. Uribe et al., *Electrochem. Soc. Proc.* **1995**, *95-23*, 284.
64. K. D. Kreuer, *J. Membr. Sci.* **2001**, *185*, 29.
65. G. Alberti, M. Casciola, L. Massinelli et al., *J. Membr. Sci.* **2001**, *185*, 73.
66. F. N. Büchi, G. G. Scherer, *J. Electrochem. Soc.* **2001**, *148*, A183.
67. X. Z. Du, J. Yu, M. Han et al., *Phys. Chem. Chem. Phys.* **2001**, *3*, 3175.
68. J. Kerres, *J. Membr. Sci.* **2001**, *185*, 3.
69. B. Gupta, F. N. Büchi, G. G. Scherer, *Solid State Ionics* **1993**, *61*, 213.
70. T. Lehtinen, G. Sundholm, S. Holmberg et al., *Electrochim. Acta* **1997**, *43*, 1881.
71. J. S. Wainright, J.-T. Wang, D. Weng et al., *J. Electrochem. Soc.* **1995**, *142*, L121.
72. S. M. Haile, D. A. Boysen, C. R. I. Chisholm et al., *Nature* **2001**, *410*, 910.
73. K. D. Kreuer, A. Fuchs, M. Ise et al., *Electrochim. Acta* **1998**, *43*, 1281.
74. J. J. Sumner, S. E. Creager, J. J. Ma et al., *J. Electrochem. Soc.* **1998**, *145*, 107.
75. P. Jannasch, *Curr. Opin. Colloid Interface Sci.* **2003**, *8*, 96.

76. F. Wang, M. Hickner, Y. S. Kim et al., *J. Membr. Sci.* **2002**, *197*, 231.
77. Y. Yang, S. Holdcroft, *Fuel Cells* **2005**, *5*, 171.
78. A. Lehmani, O. Bernard, P. Turq, *J. Stat. Phys.* **1997**, *89*, 379.
79. J. L. Fales, N. E. Vanderbrogh, P. Stroeve, in *Diaphragms, Separators, and Ion-Exchange Membranes* (Eds.: J. W. Van Zee, R. E. White, K. Kinoshita et al.), Proc. Vol. 86-13, The Electrochemical Society, Pennington, 1986, p. 179.
80. F. A. L. Dullien, *Porous Media*, Academic Press, New York, 1979.
81. Yu. M. Volfkovich, N. V. Kuleshov, and E. L. Filippov, *Sov. Electrochem.* **1980**, *16*, 1512.
82. C. Fabiani, S. Scuppa, L. Bimdi et al., *J. Electrochem. Soc.* **1983**, *130*, 583.
83. M. Eikerling, *Theoretische Modellierung der elektrophysikalischen Eigenschaften, der Struktur und Funktion von Niedertemperatur-Ionenaustauschemembranen*, Dissertation, TU, München, 1999, Jülich-Report-3717, 2000.
84. A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2001**, *105*, 9586.
85. T. A. Zawodzinski Jr., T. E. Springer, F. Uribe et al., *Solid State Ionics* **1993**, *60*, 199.
86. T. A. Zawodzinski Jr., J. Davey, J. Valerio et al., *Electrochim. Acta* **1995**, *40*, 297.
87. M. Ise, K. D. Kreuer, J. Maier, *Solid State Ionics* **1999**, *125*, 213.
88. M. Sahimi, *Flow and Transport in Porous Media and Fractured Rock*, VCH, Weinheim, 1995.
89. M. Miyahara, H. Kanda, T. Yoshioka et al., *Langmuir* **2000**, *16*, 4293.
90. W. J. Stroud, J. E. Curry, J. H. Cushman, *Langmuir* **2001**, *17*, 688.
91. D. Stauffer, A. A. Aharony, *Introduction to Percolation Theory*, revised 2nd edition, Taylor & Francis, London, 1994.
92. F. N. Büchi, J. Huslage, G. G. Scherer, PSI – Annual Report 1997, Annex V, General Energy Research, 48.
93. R. Mosdale, G. Gebel, M. Pineri, *J. Membr. Sci.* **1996**, *118*, 269.
94. F. N. Büchi, G. G. Scherer, *J. Electroanal. Chem.* **1996**, *404*, 37.
95. A. A. Kulikovsky, *J. Electrochem. Soc.* 2003, *150*, A1432.
96. M. Wakizoe, O. A. Velev, S. Srinivasan, *Electrochim. Acta* **1995**, *40*, 335.
97. K. Bolwin, E. Gülzow, D. Bevers et al., *Solid State Ionics* **1995**, *77*, 324.
98. S. Escribiano, P. Aldebert, *Solid State Ionics* **1995**, *77*, 318.
99. S. Gottesfeld, T. A. Zawodzinski Jr., *Polymer Electrolyte Fuel Cells*, in *Advances in Electrochemical Science and Engineering* (Eds.: R. C. Alkire, H. Gerischer, D. M. Kolb et al.), Wiley-VCH, Weinheim, New York, 1997, p. 195, Vol. 5.
100. T. R. Ralph, G. A. Hards, G. E. Keating et al., *J. Electrochem. Soc.* **1997**, *144*, L267.
101. M. Uchida, Y. Aoyama, N. Eda et al., *J. Electrochem. Soc.* **1995**, *142*, 463.
102. M. Uchida, Y. Aoyama, N. Eda et al., *J. Electrochem. Soc.* **1995**, *142*, 4143.
103. R. B. Bird, W. E. Stewart, E. N. Lightfoot, *Transport Phenomena*, Wiley, New York, 1960.
104. M. M. Mezedur, M. Kaviany, W. Moore, *AIChE J.* **2002**, *48*, 15.
105. Yu. A. Chizmadzev, Yu. G. Chirkov, Chapter 5, *Electrodics: Transport*, in *Comprehensive Treatise of Electrochemistry* (Eds.: J. O. M. Bockris, Yu. A. Chizmadzhev, B. E. Conway et al.), Plenum Press, New York, 1983, pp. 317–391, Vol. 6.
106. J. S. Newman, *Electrochemical Systems*, Prentice Hall, 1991.
107. M. Eikerling, A. A. Kornyshev, A. S. Ioselevich *Fuel Cells* **2004**, *4*, 131.
108. A. A. Kulikovsky, *Electrochem. Commun.* **2002**, *4*, 318.
109. A. C. Fisher, *Electrode Dynamics*, Oxford University Press, Oxford, 1996.
110. A. Parthasarathy, S. Srinivasan, A. J. Appleby et al., *J. Electrochem. Soc.* **1992**, *139*, 2856.
111. A. Damjanovic, D. B. Sepa, M. V. Vojnovic, *Electrochim. Acta* **1979**, *24*, 887.
112. A. M. Kuznetsov, J. Ulstrup, *Electron Transfer in Chemistry and Biology*, *Wiley Series in Theoretical Chemistry*, John Wiley & Sons, New York, 1998.
113. C. H. Haman, W. Vielstich, *Elektrochemie II*, Verlag Chemie, Physik-Verlag, Weinheim, 1981, p. 26.
114. O. A. Khazova, A. A. Mikhailova, A. M. Skundin et al., *Fuel Cells* **2002**, *2*, 99.
115. L. X. You, H. Liu, *Int. J. Hydrogen Energy* **2001**, *26*, 991.
116. J. Ihonen, F. Jaouen, G. Lindbergh et al., *J. Electrochem Soc.* **2002**, *149*, A448.

117. J. Xie, K. L. Moore, T. A. Zawodzinski et al., *J. Electrochem. Soc.* **2004**, *151*, A1841.
118. M. B. Isichenko, *Rev. Mod. Phys.* **1992**, *64*, 961.
119. A. S. Ioselevich, A. A. Kornyshev, *Phys. Rev. E* **2002**, *65*, 2131.
120. A. Ioselevich, A. A. Kornyshev, W. Lehnert, *Solid State Ionics* **1999**, *124*, 221.
121. R. P. Simpraga, B. E. Conway, *Electrochim. Acta* **1998**, *43*, 3045.
122. Q. Wang, M. Eikerling, D. Song et al., *J. Electrochem. Soc.* **2004**, *151*, A950.
123. Z. Xie, T. Navessin, Z. Shi et al., *J. Electrochem. Soc.* **2005**, *152*, A1171.
124. D. Song, Q. Wang, S. Liu et al., *J. Power Sources* **2004**, *126*, 104.
125. T. Navessin, S. Holdcroft, Q. Wang et al., *J. Electroanal. Chem.* **2004**, *567*, 111.
126. T. Navessin, M. Eikerling, Q. Wang et al., *J. Electrochem. Soc.* **2005**, *152*, A796.
127. F. Maillard, M. Eikerling, O. V. Cherstiouk et al., *Faraday Discuss* **2004**, *125*, 357.
128. A. S. Ioselevich, A. A. Kornyshev, *Fuel Cells* **2001**, *1*, 40.
129. Q. Wang, M. Eikerling, D. Song et al., *J. Electroanal. Chem.* **2004**, *573*, 61.
130. J. J. Macdonald (Ed.), *Impedance Spectroscopy*, Wiley, New York, 1987.
131. Z. B. Stoynov, B. M. Grafov, B. S. Savova-Stoynov et al., *Electrochemical Impedance*, Nauka, Moscow, 1991.
132. R. De Levie, in *Advances in Electrochemistry and Electrochemical Engineering* (Ed.: P. Delahay), Wiley, New York, 1967, p. 329, Vol. 6.
133. F. Gutmann, *J. Appl. Electrochem.* **1965**, *112*, 94.
134. H. Keiser, K. Beccu, M. A. Gutjahr, *Electrochim. Acta* **1976**, *21*, 539.
135. M. Keddam, C. Rakotomavo, H. Takenouti, *Electrochim. Acta* **1984**, *14*, 437.
136. N. A. Hampson, S. A. G. R. Karunathilaka, R. Leek, *J. Appl. Electrochem.* **1980**, *10*, 3.
137. J. I. Zhang, F. Wenger, M. Jerome et al., *C.R. Acad. Sci. C* **1985**, 995.
138. P. Lay, P. F. Lawrence, N. J. M. Wilkins et al., *J. Appl. Electrochem.* **1985**, *15*, 755.
139. *Proceedings of 4th International Symposium on Electrochemical Impedance Spectroscopy*, Rio de Janeiro, Brazil, 1998.
140. *Proceedings of.5th International Symposium on Electrochemical Impedance Spectroscopy*, Marileva, Italy, 2001.
141. T. E. Springer, I. D. Raistrick, *J. Electrochem. Soc.* **1989**, *136*, 1594.
142. I. D. Raistrick, *Electrochim. Acta* **1990**, *35*, 1579.
143. T. E. Springer, T. A. Zawodzinksi, M. S. Wilson et al., in *Proceedings of the First International Symposium on Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gottesfeld, G. Halpert, A. Landgrebe), Proc. 95-23, The Electrochemical Society, Pennington, NJ, 1995.
144. M. Eikerling, A. A. Kornyshev, *J. Electroanal. Chem.* **1999**, *475*, 107.
145. A. J. Bard, L. R. Faulkner, Chapter 12, *Electrochemical Methods*, Wiley, New York, 1980, p. 501.
146. H. Natter, R. Hempelmann, *Electrochim. Acta* **2003**, *49*, 51.
147. M. K. Debe, in *Handbook of Fuel Cells: Fundamentals, Technology, and Applications*, (Eds.: W. Vielstich, A. Lamm et al.), John Wiley & Sons, New York, 2002, Vol. 3, Ch. 45, p. 576.
148. J. Feders, *Fractals*, Plenum, New York, 1988.
149. B. B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman, New York, 1982.
150. M. S. Wilson, *Thin Film Electrode*, US patent #5211984, May 18, 1993.
151. P. Zelenay, S. C. Thomas, S. Gottesfeld, *Direct Methanol Fuel Cells: Recent Progress in Fuel Efficiency, Cell Performance and Performance Stability*, in *Proton Conducting Membrane Fuel Cells (Second International Symposium)* (Eds.: S. Gottesfeld, T. F. Fuller, G. Halpert), Proc. Vol. 98-27 Electrochemical Society, Pennington, 1999, pp. 300–315.
152. S. C. Thomas, X. Ren, S. Gottesfeld and P. Zelenay, ''Direct Methanol Fuel Cells: Progress in Cell Performance and Cathode Research,'' *Electrochim. Acta* **2002**, *47*, 3741.
153. J. Mergel, D. Stolten, H. Schmitz et al., *Presented at 2002 Fuel Cell Seminar*, Palm Springs, Nov. 18 – 21, 2002.
154. K. Scott, W. Taama, J. Cruichshank, *J. Appl. Electrochem.* **1998**, *28*, 289.
155. S. F. Baxter, V. S. Battaglia, R. E. White, *J. Electrochem. Soc.* **1999**, *146*, 437.
156. K. Scott, P. Argyropoulos, K. Sundmacher, *J. Electroanal. Chem.* **2000**, *477*, 97.
157. H. Dohle, J. Divisek, R. Jung, *J. Power Sources* **2000**, *86*, 469.
158. G. Murgia, L. Pisani, M. Valentini et al., *J. Electrochem. Soc.* **2002**, *149*, A31.

159. J. S. Yi, T. V. Nguyen, *J. Electrochem. Soc.* **1998**, *145*, 1149.
160. V. Gurau, H. Liu, S. Kakac, *AIChE J.* **1998**, *44*, 2410.
161. K. Dannenberg, P. Ekdunge, G. Lindbergh, *J. Appl. Electrochem.* **2000**, *30*, 1377.
162. S. Um, C.-Y. Wang, K. S. Chen, *J. Electrochem. Soc.* **2000**, *147*, 4485.
163. P. Futerko, I.-M. Hsing, *Electrochim. Acta* **2000**, *45*, 1741.
164. I.-M. Hsing, P. Futerko, *Chem. Eng. Sci.* **2000**, *55*, 4209.
165. Z. H. Wang, C. Y. Wang, K. S. Chen, *J. Power Sources* **2001**, *94*, 40.
166. T. E. Springer, T. Rockward, T. A. Zawodzinski et al., *J. Electrochem. Soc.* **2001**, *148*, A11.
167. A. Kazim, H. T. Liu, P. Forges, *J. Appl. Electrochem.* **1999**, *29*, 1409.
168. J. S. Yi, T. V. Nguyen, *J. Electrochem. Soc.* **1999**, *146*, 38.
169. A. A. Kulikovsky, *J. Appl. Electrochem.* **2000**, *30*, 1005.
170. W. He, J. S. Yi, T. V. Nguyen, *AIChE J.* **2000**, *46*, 2053.
171. D. Natarajan, T. V. Nguyen, *J. Electrochem. Soc.* **2001**, *148*, A1324.
172. S. Dutta, S. Shimpalee, J. W. Van Zee, *J. Appl. Electrochem.* **2000**, *30*, 135.
173. S. Shimpalee, S. Dutta, *Num. Heat Transfer* **2000**, *38*, 111.
174. J. Mosig, *Dreidimensionale mathematische Modellierung einer Brennstoffzelle unit Protonen-Austausche Membran (PEMFC)* PhD thesis, Forschungszentrum Jülich, Jül-3480 D82, ISSN 0944–2952, 1997.
175. J. Divisek, J. Mosig, B. Steffen, *Proton Exchange Membrane Fuel Cell Model*, in *Electrochemical Engineering and Energy* (Eds.: F. Lapicque, A. Stork, A. A. Wragg), Plenum Press, New York, 1995, pp. 187–196.
176. J. Divisek, private communication, 1998.
177. A. A. Kulikovsky, *Electrochem. Commun.* **2001**, *3*, 460.
178. A. A. Kulikovsky, *Fuel Cells* **2001**, *1*, 162.
179. V. S. Bagotzky, Yu. B. Vasilyev, *Electrochim. Acta* **1964**, *9*, 869.
180. C. Lamy, J.-M. Leger, *J. Chem. Phys.* **1991**, *88*, 1649.
181. S. J. Ridge, R. I. White, Y. Tsou et al., *J. Electrochem. Soc.* **1989**, *136*, 1902.
182. A. S. Arićo, V. Antonicci, V. Alderucci et al., *J. Appl. Electrochem.* **1993**, *23*, 1107.
183. A. Parthasarathy, C. R. Martin, S. Srinivasan, *J. Electrochem. Soc.* **1991**, *138*, 916.
184. P. S. Kauranen, E. Skou, *J. Electroanal. Chem.* **1996**, *408*, 189.
185. P. S. Kauranen, E. Skou, J. Munk, *J. Electroanal. Chem.* **1996**, *404*, 1.
186. T. J. Schmidt, H. A. Gasteiger, R. J. Behm, *Electrochem. Commun.* **1999**, *1*, 1.
187. A. Wieckowski, E. Savinova, C. Vayenas (Eds.), *Catalysis and Electrocatalysis at Nanoparticle Surfaces*, Marcel Dekker, New York, 2003.
188. O. V. Cherstiouk, P. A. Simonov, E. R. Savinova, *Electrochim. Acta* **2003**, *48*, 385.
189. V. P. Zhdanov, B. Kasemo, *Surf. Sci. Rep.* **2000**, *39*, 25–104.
190. W. H. Press, B. P. Flannery, S. A. Teukolsky et al., *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1993.
191. D. N. de G. Allen, R. V. Southwell, *Q. J. Mech. Appl. Math.* **1955**, *8*, 129.
192. A. A. Kulikovsky, *Electrochem. Commun.* **2001**, *3*, 572.
193. K. Scott, P. Argyropoulos, P. Yiannopoulos et al., *J. Appl. Electrochem.* **2001**, *31*, 823.
194. A. A. Kornyshev, A. A. Kulikovsky, *Electrochim. Acta* **2001**, *45*, 4389.
195. H. Dohle, A. A. Kornyshev, A. A. Kulikovsky et al., *Electrochem. Commun.* **2001**, *3*, 73.
196. A. A. Kulikovsky *Electrochim. Acta* **2004**, *49*, 617.
197. D. J. L. Brett, S. Atkins, N. P. Brandon et al., *Electrochem. Commun.* **2001**, *3*, 628.
198. A. A. Kulikovsky *Electrochem. Commun.* **2002**, *4*, 845.
199. A. A. Kulikovsky, A. Kucernak, A. A. Kornyshev *Electrochim. Acta* **2005**, *50*, 1323.
200. M. M. Mench, C. Y. Wang, M. Ishikawa. *J. Electrochem. Soc.* **2003**, *150*, 1052.
201. A. A. Kulikovsky, H. Scharmann, K. Wipperman. *Electrochem. Commun* **2004**, *6*, 729.
202. A. S. Arićo, P. Cretì, V. Baglio et al., *J. Power Sources* **2000**, *91*, 202.
203. A. S. Arićo, S. Srinivasan, V. Antonucci, *Fuel Cells* **2001**, *1*, 133.
204. M. S. Wilson, T. E. Springer, J. R. Davey et al., in *Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gottesfeld, G. Halpert, A. Langrebe), Proc. Ser. PV 95–23, The Electrochemical Society, Pennington, 1995, p. 115.

205. T. V. Nguyen, *J. Electrochem. Soc.* **1996**, *143*, L103.
206. D. L. Wood, J. S. Yi, T. V. Nguyen III, *Electrochim. Acta* **1998**, *43*, 3795.
207. A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2000**, *104*, 4471.
208. A. Vishnyakov, A. V. Neimark, *J. Phys. Chem. B* **2001**, *105*, 7830.
209. J. Ennari, M. Elomaa, F. Sundholm, *Polymer* **1999**, *40*, 5035.
210. K. A. Friedrich, K. P. Geyzers, A. J. Dickinson et al., *J. Electroanal. Chem.* **2002**, *524*, 261.
211. K. A. Mauritz and R. B. Moore, *Chem Rev.* **2004**, *104*, 4535.
212. K. D. Kreuer, S. J. Paddison, E. Spohr et al., *Chem Rev.* yr2004, *104*, 4637.
213. A. Z. Weber and J. Newman, *Chem Rev.* **2004**, *104*, 4679.
214. N. W. Deluca and Y. A. Elabd, *J. Polym. Sci. B, Polym. Phys.* **2006**, *44*, 2201.
215. Q. F. Li, R. H. He, J. O. Jensen et al., *Chem. Mater.* **2003**, *15*, 4896.
216. B. Smitha, S. Sridhar, and A. A. Khan, *J. Membrane Sci.* **2005**, *259*, 10.
217. A. Vishnyakov and A. V. Neimark, *J. Phys. Chem.* **2001**, *105*, 9586.
218. S. S. Jang, V. Molinero, T. Çağin et al., *J. Phys. Chem. B* **2004**, *108*, 3149.
219. D. Seeliger, C. Hartnig, and E. Spohr, *Electrochim. Acta* **2005** *50*, 4234.
220. J. L. Matthews, E. K. Lada, L. M. Weiland et al. *Smart Mat. Struct.* **2006**, *15*, 187.
221. (a) P. G. Khalatur, S. K. Talitskikh, and A. R.Khokhlov, *Macromol. Theory Sim.* **2002**, *11*, 566–586. (b) D. Y. Galperin and A. R. Khokhlov, *Macromol. Theory Simul.* **2006**, *15*, 137.
222. J. T. Wescott, Y. Qi, L. Subramanian et al., *J. Chem. Phys.* **2006**, *124*, 134702.
223. (a) S. J. Paddison and J. A. Elliott, *J. Phys. Chem. A* **2005**, *109*, 7583; S. J. Paddison, *Ann. Rev. Mater. Res.* **2003**, *33*, 289.
224. M. K. Petersen, F. Wang, N. P. Blake et al., *J. Phys. Chem. B* **2005**, *109*, 3727–3730.
225. A. Roudgar, S. P. Narasimachary, and M. Eikerling, *J. Phys. Chem. B* in press.
226. M. Schuster, T. Rager, A. Noda et al., *J. Fuel Cells* **2005**, *5*, 355.
227. J. Teissié, M. Prats, P. Soucaille et al., *Proc. Natl. Acad. Sci.* **1985**, *82*, 3217.
228. V. B. P. Leite, A. Cavalli, and O. N. Oliveira, Jr., *Phys. Rev. E* **1998**, *57*, 6835.
229. D. Bazeia, V. B. P. Leite, B. H. B. Lima et al. *Chem. Phys. Lett.* **2001**, *340*, 205.
230. A. M. Smondyrev, and G. A. Voth, *Biophys. J.* **2002**, *82*, 1460.
231. S. Serowy, S. M. Saparov, Y. N. Antonenko et al., *Biophys. J.* **2003**, *84*, 1031.
232. A. Y. Mulkidjanian, J. Heberle, and D. A. Cheerepanov, *Biophys. Biochem. Acta* **2006** (in press).
233. C. Hartnig, P. Vassilev, and M. T. M. Koper, *Electrochim. Acta* **2003**, *48*, 3751.
234. A. Wieckowski, E. R. Savinova and C. G. Vayenas, (Eds.), *Catalysis and Electrocatalysis at Nanoparticle Surfaces*, Marcel Dekker, Inc, New York, 2003, pp.970.
235. J. Greeley, J. K. Nørskov, and M. Mavrikakis, *Ann. Rev. Phys. Chem.* **2002**, *53*, 319.
236. M. Neurock, *J. Catalysis* **2003**, *216*, 73.
237. J. Meier, J. Schiotz, P. Liu et al., *Chem. Phys. Lett.* **2004**, *390* 440.
238. V. Stamenkovic, B. S. Mun, K. J. J. Mayrhofer et al., *Ang. Chem. Int. Ed.* **2006**, *45*, 2897.
239. B. Andreaus, F. Maillard, J. B. Kocylo et al., *J. Phys. Chem. B* **2006**, in press.
240. Q. Wang, M. Eikerling, D. Song et al., *J. Electroanal. Chem.* **2004**, *573*, 61.
241. M. Eikerling, *J. Electrochem. Soc.* **2006**, *153*, E58.
242. A. Havranek, K. Wippermann, *J. Electroanal. Chem.* **2004**, *567*, 305.
243. J. Stumper, H. Haas, and A. Granados, *J. Electrochem. Soc.* **2005**, *152*, A837.
244. J. P. Meyers and J. Newman *J. Electrochem. Soc.* **2002**, *149*, A718.
245. A. A. Kulikovsky, *Electrochem. Comm.* **2005**, *7*, 969.
246. A. A. Kulikovsky, *Electrochem. Comm.* **2004**, *6*, 969.
247. S. M. Senn, D. Poulikakos *Electrochem. Comm.* **2005**, *7*, 773.
248. A. A. Kulikovsky *J. Electrochem. Soc.* **2005**, *152*, A1121.
249. S. M. Senn, D. Poulikakos *Int. J. Heat and Mass Transf.*, **2006**, *49*, 1516.
250. A. A. Kulikovsky, H. Scharmann, K. Wippermann, *Electrochem. Comm.* **2004**, *6*, 75.
251. A. A. Kulikovsky, H. Scharmann, K. Wippermann *Electrochem. Comm.* **2004**, *6*, 729.
252. I. Nazarow and K. Promislow, *Chem. Eng. Sci.* **2006**, *61*, 3198.
253. A. A. Kulikovsky, *Electrochem. Comm.* **2005**, *7*, 237.

254. A. A. Kulikovsky *Electrochim. Acta* **2005**, **2006**, *51*, 2003.
255. A. A. Kulikovsky, *Electrochem. Comm.* **2004**, *6*, 1259.
256. Q. Ye, T. S. Zhao, H. Yang et al., *Electrochem. Solid – State Lett.* **2005**, *8*, A52.
257. M. Mueller, H. Dohle, A. A. Kulikovsky, *Electrochem. Solid – State Lett.* **2005**, *9*, L7.
258. Q. Ye, T. S. Zhao, *J. Electrochem. Soc.* **2005**, *152*, A2238.
259. A. A. Kulikovsky, K. Klafki, K. Wippermann, *Electrochem. Comm.* **2005**, *7*, 394.
260. A. A. Kulikovsky, H. Schmitz, K. Wippermann et al., *Electrochem. Comm.* **2006**, *8*, 754.
261. G. S. Kim, J. St-Pierre, K. Promislow et al., *J.Power Sources* **2005**, *152*, 210.
262. A. A. Kulikovsky J. Power Sources 2006, doi: 10.1016/j.jpowsour.2006.01.058.
263. P. Berg, A. Caglar, K. Promislow et al., *IMA J. Appl. Math.* **2005**, *71*, 241.
264. A. A. Kulikovsky, *J.Electrochem. Soc.* **2006**, *153*, A1672.
265. A. A. Kulikovsky, *Electrochem. Comm.* **2006**, *8*, 1225.
266. S. Shimpalee, S. Greenway, D. Spuckler et al., *J. Power Sources* **2004**, *135*, 79.
267. H. Meng, C.-Y. Wang, *J. Electrochem. Soc.* **2005**, *152*, A1723.
268. E. Birgersson, J. Nordlund, H. Ekstroem et al., *J. Electrochem. Soc.* **2003**, *150*, A1368.
269. E. Birgersson, J. Nordlund, M. Vynnycky, et al., *J. Electrochem. Soc.* **2004**, *151*, A2157.
270. Ulf Bossel The Birth of the Fuel Cell, European Fuel Cell Forum, 2000. ISBN 3-905592-09-6. Offset – and Dissertations-Druck Juergen Kinzel, Goettingen, Germany.

8.3 Polymer Electrolyte and Direct Methanol Fuel Cells

Shimshon Gottesfeld
MTI Microfuel Cells, Albany, New York, USA and Fuel Cell Consulting, LLC, Niskayuna, LLC, Niskayuna, NY, USA

8.3.1 Introduction

This chapter covers electrochemical science and engineering aspects of polymer electrolyte hydrogen/air fuel cell (PEFC) and direct methanol/air fuel cell (DMFC) technologies. These two technologies underwent exciting developments during the last two decades, enabled by a wide range of efforts worldwide to advance from R&D to prototype building and testing of PEFC and DMFC power systems targetting market entry of early commercial products. This introduction to the chapter provides a brief overview of key challenges in these areas of electrochemical technology, together with the R&D work done to date to advance the technology. Technology elements, briefly covered in the introduction, reflect those covered in the body of the chapter: proton-conducting polymeric membranes, PEFC membrane/electrode assemblies (MEAs) and PEFC stacks, electrocatalysis and transport in the PEFC anode and PEFC cathode, and last, but not least, membrane electrolytes and electrode processes in DMFCs and the potential of small-scale DMFC systems to serve as next-generation batteries for handsets and laptops.

The PEFC has been one of the fastest growing fields of research, development, and engineering in the area of applied electrochemistry over the last 20 years (1985–2005). The growing interest in this type of fuel cell has been the result of the ever-growing search for cleaner and more efficient energy-conversion processes in major areas of fuel and energy consumption, including primarily automotive transportation and generation of electric power. As an electrochemical, fuel/oxygen cell that converts the chemical energy of the fuel directly to electric energy, this type of power source has some intrinsic attractive features. One such feature is the high efficiency, η_{theor}, theoretically achievable at temperatures under 100 °C in direct conversion of the chemical energy of a fuel into electric energy. η_{theor} is calculated as:

$$\eta_{\text{cell, theor}} = \Delta G_{\text{cell process}} / \Delta H_{\text{cell process}} \quad (1)$$

where $\Delta G_{\text{cell process}}$ is the Gibbs free energy change and $\Delta H_{\text{cell process}}$ the enthalpy change in the cell process (the fuel "heating value") at the relevant cell temperature and activities of reactants and products. This theoretical efficiency exceeds 80% for a fuel cell directly converting hydrogen or methanol fuel to electric power at operation temperatures typical for a PEFC, that is, 30–100 °C. Given that the practical energy-conversion efficiency of heat engines is typically below 30% and approaches 50% only for gas turbines operating at much higher temperatures, a fuel-cell power source of heat engine level power density, operating near ambient temperatures and converting the chemical to electric energy at 50% efficiency is an attractive alternative. The argument for low temperature power source operation is strong in transportation applications, where fast "cold start-up" is of high premium. So is elimination of energy losses associated with repeated heating of power sources operating at much higher temperatures, particularly when several on/off cycles per day are involved.

Encyclopedia of Electrochemistry.Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.ISBN: 978-3-527-30397-7

Beyond the high conversion efficiency, the other very significant promise of fuel cells as new technology for transportation and electric power generation is the environmental benefits. These include lowering, or eliminating altogether of hazardous tail pipe emissions and, in addition, reducing greenhouse gas emissions. Nitrogen oxide (NO_x), a major pollutant emitted by internal combustion engines is not generated by a fuel-cell power source operating at 80 °C or below. Other hazardous emissions, for example, carbon monoxide, are also drastically lowered in a fuel-cell-powered vehicle and are, in fact, brought down to zero (''zero-emission vehicle'') when hydrogen is used as fuel for the PEFC, and water vapor becomes the only exhaust from the power system.

Whereas transportation and off-grid power generation have been highlighted to date as the central potential applications of PEFCs, applications of PEFCs and, particularly, of DMFCs for portable power sources may actually provide the first market entry for these low-temperature fuel-cell technologies. The latter type of application has generated a significant increase in interest and activity during the last 5 years, driven by the clear need in the consumer electronics industry, as well as in the military sector, for power sources of energy density exceeding that of advanced Li batteries. Such advanced power sources are required to power present and, particularly, next-generation multifunctional handheld devices, to achieve longer use times. The projection of faster market entry in this sector of the technology is based on the evident potential to significantly exceed the energy density of present-day portable power sources and, at the same time, compete favorably on cost with the incumbent battery technology.

The PEFC was developed originally in the 1960s and represented a new and different type of low-temperature fuel-cell technology. When compared with the two low-temperature fuel-cell technologies introduced earlier, the alkaline fuel cell (AFC) and the phosphoric acid fuel cell (PAFC), the outstanding feature of the PEFC is a polymer membrane electrolyte of high protonic conductivity requiring only deionized water to facilitate protonic conductivity in the polymer. Consequently, a significant advantage of the PEFC over other, low-temperature fuel cells is the elimination of problems typically associated with a liquid electrolyte, including electrolyte migration within and electrolyte loss from the fuel cell. It could be argued that these problems have been replaced by the need of effective ''water management'' to achieve optimized water level and distribution in an operating PEFC, but, while water management presents some hurdles in PEFC stack and system design, the PEFC system remains significantly more manufacturer and user friendly by avoiding the handling and system management of corrosive liquid electrolytes.

In the development of fuel-cell technology based on this unique polymer electrolyte, special chapters in electrochemical science and engineering have emerged, addressing the fuel-cell ionomeric membrane itself and the optimized fabrication of MEAs. The invention of Nafion, a poly(perfluorosulfonic acid) (poly(PFSA)) at DuPont in the 1960s, was, in fact, a key (if not *the* key) milestone in the development of PEFC technology. The chemical and mechanical properties of such poly(PFSA) extruded membranes, which are based on a perfluorocarbon backbone, enabled to achieve stable materials properties and, consequently,

stable PEFC performance over thousands of hours. Under constant load conditions purely proton conductor ($t_{H^+} = 1$), the protonic mobility in the polymer membrane electrolyte, is a strong function of water content and, consequently, proper measurement and modeling of membrane transport properties, and their role in effective water management in an operating PEFC or DMFC, has contributed significantly to technology advancements.

The (pseudo) solid nature of the electrolyte in the PEFC system presented a challenge to researchers and technologists in this field as they came to optimize the electrode/membrane electrolyte interface to maximize catalyst utilization. With distilled water as the only liquid in the cell, ionic mobility is confined to the hydrated ionomer (ion-containing polymer). For this, an electrode/electrolyte interface had to be designed and optimized for a process where the reactant is gaseous (O_2 or H_2) and the overall available active surface area is defined by carbon-supported metal, or metal alloy catalyst particles typically 2–5 nm in diameter, that have to maintain good contact with the hydrated ionomer. Work during the last 20 years has generated a good understanding and a detailed physicochemical description of the mixed conductor layer containing the dispersed catalyst particles, that is required to maximize catalyst utilization in the PEFC electrode. This resulted in continued optimization of the structure and composition of such catalyst layers and their bonding to the membrane surface. The latter specific area of activity is a very important facet of the science and engineering of PEFCs and DMFCs, paving the way to cost-effective technology.

When it comes to the intrinsic activity of the metal electrocatalyst itself, both the PEFC and DMFC technologies can be described as ''strongly challenged''. An air (oxygen) electrode operating in an acid electrolyte presents a well-documented challenge, which is further exacerbated if interfacial water activity falls below unity – a potential problem in the case of the membrane proton conductor. Power densities exceeding 1 kW L^{-1}, achieved in hydrogen/air PEFC stacks, are therefore indication for the high quality of PEFC air electrodes developed on the basis of catalyst layers of optimized compositions and proper cell humidification. With such electrochemical and materials science tools, the intrinsically low oxygen reduction reaction (ORR) activity of best-available catalysts can be remedied by effectively utilized catalysts of high surface area. The electrocatalysis challenge on the fuel (anode) side of the PEFC is of different nature, but not much less demanding. Achieving satisfactory ''CO tolerance'', as demanded by hydrogen feed streams derived from carbonaceous fuels, has been one of the most intriguing aspects of the electrochemistry of the PEFC and the effective chemical and/or electrochemical approaches developed to provide good CO tolerance are key technology enablers. It is, however, the DMFC that presents the higher electrocatalysis challenge in the fuel (anode) side of the cell. Polymer electrolyte DMFC technology is based on direct and complete electrooxidation of a carbonaceous liquid fuel, that has to be maintained at high rate for thousands of hours at temperature close to ambient. This is highly demanding catalytically; however, it deserves the effort as these fuel cells that convert directly liquid fuel at low temperature are of high value, particularly as man-portable power sources. Various electrochemical and power system aspects of

polymer electrolyte DMFCs are covered in the last part of this chapter.

Finally, the fuel-cell *stack* and overall PEFC, or the DMFC *power system*, involve many elements beyond the electrochemical cell itself. The PEFC stack requires effective design and fabrication to achieve uniform current distribution across large cell areas (up to 1 ft^2), as well as uniform performance of many individual cells connected electrically in series and having typically parallel fuel and air feed streams. The stack also requires effective heat management to dissipate a heat load typically larger than the electric power generated, under quite limited temperature gradients available between the stack and ambient air. PEFC stack engineering aspects have strongly advanced hand in hand with the advancements in electrochemical and materials science aspects of the technology, with equally important contributions made from both ends to fabrication of high performance, compact PEFC and DMFC stacks.

This field of science and technology was reviewed during the last decade in several papers and books including specific chapters devoted to PEFCs. Two earlier contributions were devoted to a general review of PEFCs [1] and to the potential application of PEFCs and DMFCs for portable power sources [2]. A recent book [3] covers engineering and systems aspects of fuel cells, including significant parts devoted to PEFCs. For Electrochemical Society Symposia Proceeding Volumes cover symposia devoted specifically to the science and technology of polymer electrolyte PEFCs and DMFCs [4]. Finally, a comprehensive source of information on fuel-cell science and technology was provided in the form of a four-volume ''Handbook of Fuel Cells'' [5], published in 2003. The reader is referred to the above for further relevant material and insight.

8.3.2 The Polymer Electrolyte and Direct Methanol Fuel Cells: Brief Overview of Key Components and Features of Cells and Stacks

Figure 1 provides a schematic presentation of the cross section of a single PEFC. This scheme is used to discuss key materials and processes in the PEFCs. The ''heart'' of the cell, which is magnified in the scheme, is the so-called MEA. In its simplest form, the ''electrode'' component of the MEA consists of a thin film (typically 5–10 μm thick) containing a dispersed Pt catalyst. This catalyst layer is in good contact with the ionomeric membrane (the central slab in the scheme in Fig. 1), which serves as electrolyte and fuel/oxidant separator in this cell. The membrane electrolyte ranges in thickness between 25 and 175-μm thickness, with the thinner values usually associated with high-power-density PEFCs and the larger values more typical for DMFCs. The ''3-layer'' combination of a polymeric membrane with two catalyst layers has been referred to sometimes as catalyst-coated membrane (''CCM''), with the term MEA reserved for the ''5-layer'' combination including the two gas-diffuser layers (GDLs) (also called *backing layers*) in immediate contact with the catalyzed membrane (see Fig. 1). The latter are made of wet-proofed porous carbon paper, or carbon cloth, typically 100–300-μm thick. Wet proofing is done by coating the fibers of the carton paper/cloth with polytetrafluoroethylene (PTFE), typically employing for this purpose an aqueous colloidal suspension of PTFE.

In other modes of PEFC fabrication, the catalyst layer is applied to the porous

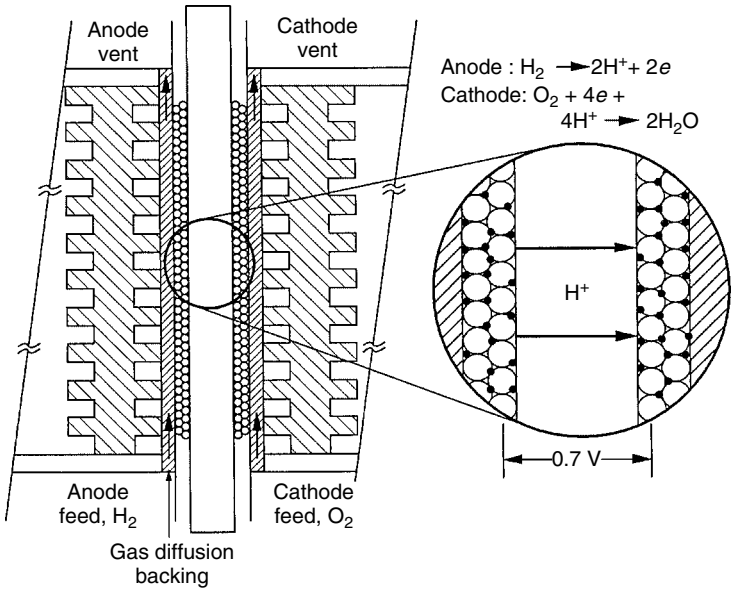


Fig. 1 Schematic presentation of a cross section of a single PEFC cell. The cell (left) consists of a proton-conducting membrane catalyzed on each side by a thin film of mixed electronic and protonic conductivity, containing uniformly distributed nanometer size catalyst particles. The gaseous reactants access the catalyst layer through a wet-proofed gas-diffusion backing layer (GDL) attached to each of the catalyst layer outer surfaces. The gaseous reactants pass along the active area of the cell (corresponding to a plane perpendicular to the page) via flow fields that are machined or molded in the current collectors. The current collectors/gas distributors are shown in the scheme as the outermost elements of the unit cell. They become "bipolar plates" in the PEFC stack, each such plate having a flow field for fuel on one side and for air on the other side of an element of high electronic conductivity and negligible gas permeability. The so-called MEA is described schematically in the magnified drawing on the right and corresponds to the most common case of a PEFC catalyst layer consisting of carbon-supported platinum particles, with a diameter of typically 10 nm (carbon) and 2–3 nm (Pt); such Pt/C particles are mixed with proton-conducting medium of recast ionomer and applied to the membrane surface as a catalyst layer typically 1–10 μm thick. Completing the MEA are the two GDLs (hatched) pressed onto the outer surface of each of the catalyst layers.

carbon backing-layer and this catalyzed carbon paper is subsequently hot pressed onto the membrane. For this mode of fabrication, the term *electrode* usually refers to the carbon paper (or carbon cloth) with the catalyst layer on one of its surfaces – the surface to be bonded in turn to the membrane. Irrespective of how the catalyst layer is bonded, first to the membrane or first to the porous carbon backing, the complete MEA of a PEFC has a cross-sectional structure which, when visually scanned from the center out in both directions, involves three adjacent layers: ionomeric membrane/catalyst layer/porous wet-proofed carbon backing (Fig. 1). The single-cell

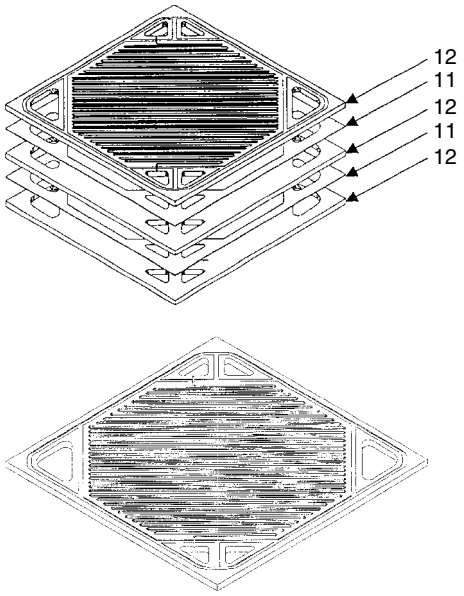


Fig. 2 Generic scheme of two adjacent cells in a bipolar stack, highlighting the general form and functions of the bipolar plate. In view is the major plane of the bipolar plate, showing a serpentine flow field designed to evenly distribute the reactant – fuel or air – across the active surface of an adjacent MEA surface (MEAs are designated in the drawing by "11" and bipolar plates by "12"). Each bipolar plate has a flow field of such general type on each of its major sides, with fuel flowing along one side of each plate and air on the other side of the same plate, each plate thereby feeding the cathode side of one MEA and the anode side of the next MEA in the stack. To reach each flow field of an individual cell, the fuel and air both flow along the axial dimension of the stack through manifolds defined by two sets of openings in the plates (top and bottom divided triangular features in the scheme). Access from an inlet manifold to the flow field and back to an outlet manifold is provided on one side of each plate for fuel, and on its opposite side for air. (The additional openings shown on the horizontal edges of the plates stand for optional coolant inlet and outlet manifolds, the use of which would require to add coolant flow fields not shown in this scheme.)

basic structure is completed by current collector plates with flow fields that could be machined on the major surfaces of the plate, or introduced as part of the fabrication of molded current collector plates. Such flow fields are required for effective distribution of reactant gases across the surface of the PEFC electrode that could be as large as 1000 cm^2.

A current collector with flow fields on each of the two major surfaces and zero gas permeability across the thickness dimension becomes the "bipolar plate" in the fuel-cell stack. The MEA is typically framed to seal the periphery and confine flow to the designed flow fields. As the generic scheme in Fig. 2 shows, matched openings in the corners of the MEA frame and the bipolar plates define separate manifolds for fuel and air supply along the length dimension of the stack. The fuel and air feed streams enter the anode and cathode of every single cell in the stack from the respective manifolds through a corner inlet point. The exhaust stream leaves through a diagonally opposite outlet point. Air and fuel could, in principle, be designed for parallel or for counterflow along the active surface of the MEA.

Returning to the MEA, the magnified part in Fig. 1 highlights the components of this central element of the PEFC, which consists of a proton-conducting membrane electrolyte with a composite catalyst layer adjacent to each of its surfaces. The scheme shows the catalyst layer as Pt (small circles) supported on carbon (larger circles). This type of supported catalyst has been used in most PEFCs developed in recent years for either transportation or stationary power generation applications. The Pt/C powder, with the weight percent of Pt ranging typically between 20 and 60, is prepared mostly by procedures based on controlled reduction of Pt complexes from aqueous solution to form metal particles of the order of 2 nm in diameter, supported onto the carbon and further stabilized through established principles of colloid chemistry. For use in a PEFC or DMFC, the Pt/C powder has to be effectively intermixed with recast ionomer to provide sufficient ionic conductivity within the catalyst layer. The catalyst layer can hence be described as a Pt/C//ionomer composite, where each of the three components is uniformly distributed within the volume of the layer. The scheme in Fig. 1 highlights PEFC catalyst layers employing Pt/C; however, in the case of DMFCs, ''blacks'' of Pt or Pt alloys have to date been used as frequently as carbon-supported catalysts, both at the anode, where PtRu alloy catalysts are the common choice, and on the DMFC cathode for which use of Pt black catalysts has been reported. Reasons for the preference of one form of dispersed catalyst against the other are discussed further in Sect. 8.3.5.

The proton-conducting polymeric membrane is the unique element of the PEFC, as is indeed reflected by any of the three names used for this type of cell: PEM fuel cell = polymer electrolyte membrane fuel cell, SPE™ fuel cell = solid polymer electrolyte fuel cell, or PEFC = polymer electrolyte fuel cell. In this chapter, we have preferred the latter acronym. The membrane material commonly employed in most recent PEFC technology developments is poly(PFSA). Nafion™ made by DuPont is the best-known material of this type. Similar materials are produced by Asahi Chemical and Asahi Glass (Japan), as well as by 3M (United States). W.L. Gore developed and provided a composite membrane for PEFCs, with expanded PTFE as backbone and poly(PFSA) as the active material. These structures are aimed at providing improved lateral dimensional stability, particularly to avoid excessive lateral dimensional changes during membrane drying/rehydration cycles. The combined chemical and physical properties of poly(PFSA) membranes give them significant advantages as membrane electrolyte for PEFCs. These membranes exhibit superior long-term chemical stability under both oxidative and reductive environments owing to their Teflon-like molecular backbone. The specific protonic conductivity achieved in well-humidified membranes of this type is of the order of 0.1 S cm^{-1} at typical cell operation temperatures, which translates, for example, to an areal resistance of 0.05 Ω cm^2 for a membrane of 50-μm thickness. These thin membranes are, at the same time, effective gas separators: the permeability of both oxygen and hydrogen through the membrane is of the order of 10^{-11}–10^{-10} mol cm^{-1} s^{-1} atm^{-1}, which translates to a gas crossover equivalent current density of 1–10 mA cm^{-2} through a membrane in an operating fuel cell. This ''leakage current'' is at or below 1% of

the operating current of a PEFC – typically around 1 A cm^{-2}.

The two important drawbacks of PFSA membranes are the limited range of temperatures in which they can be effectively employed and their high cost at present. The first limitation typically forces operation of PEFCs at temperatures below 100 °C, although some increase in the temperature of operation, e.g., to 120 °C, is possible at the expense of operation under pressurized steam. The upper limit on the temperature is dictated by the need of effective humidification of the membrane, a prerequisite for maximizing membrane protonic conductivity. Effective humidification requires, in turn, operation with poly(PFSA) membranes in a temperature/pressure domain corresponding to a dual-phase (liquid/vapor) water system because complete humidification requires the membrane to be in contact with liquid water. The need of liquid water in contact with the membrane to maximize cell performance also results in the need of subsystems adjacent to the stack, which ensure sufficient water content in the stack. This can be done, for example, by returning much of the water exhausting the stack cathode back to the stack using a water vapor exchanger (''enthalpy exchanger'') in which the incoming airstream flows counter to the cathode exhaust stream before entering the cathode inlet, the two streams being separated in the exchanger by a membrane of high water mobility, for example, a thin Nafion™ membrane [6]. As to the cost of the poly(PFSA) membrane, it has been repeatedly projected that the present cost of 500–1000 $/$m^2$, corresponding to 50–100 $/kW generated in the PEFC, should come down substantially as the market for the membrane increases significantly, particularly under a scenario of large-scale application in electric vehicles.

Adjacent to the ionomeric membrane on both sides are the thin films containing the catalyst (Fig. 1), usually referred to as *catalyst layers*. These catalyst layers could be based on Pt or Pt alloy ''black'' of typical mass loadings of 1–4 mg Pt/cm^2 of membrane area, as used till date in some polymer electrolyte DMFCs, or could be composites of carbon-supported Pt and recast ionomer of lower Pt loading of 0.1–0.4 mg Pt/cm^2 on each electrode, as developed for terrestrial transportation applications of PEFCs. The choice of carbon-supported versus unsupported catalyst (catalyst ''black'') depends on the intrinsic activity of the catalyst in the required electrode reaction and on cost considerations. Supported catalysts are more highly dispersed and, therefore, provide higher surface area per unit mass, or per unit cost of the precious metal catalyst. Carbon-supported Pt can typically provide catalyst surface areas of the order of 150–600 cm^2 Pt/cm^2 of membrane area, in a catalyst layer that is typically 10 μm thick or below, Pt mass loading of 0.1–0.4 mg Pt/cm^2 of the membrane. Such Pt mass loading and overall catalyst surface area per cm^2 of the membrane are typical in advanced MEA technology developed first in the 1990s for hydrogen/air PEFCs [7–13], targeting highly cost-sensitive applications, primarily in transportation. In spite of the lower dispersion achievable, unsupported catalysts (catalyst ''blacks'') are still preferred in applications where the catalyst surface area required per square centimeter of membrane cross-sectional area is 1000 cm^2 or higher, that is, an order of magnitude higher than that typically used in advanced CCM technology for PEFCs. This requirement of very high overall catalyst surface area arises because

of intrinsically slow interfacial kinetics and lower temperature of cell operation particularly when both constraints apply, as, for example, in the case of polymer electrolyte DMFCs for portable power applications. In the latter case, ''packaging'' of the carbon-supported catalyst to achieve as much as 1000 cm^2 Pt/cm^2 of membrane area would result in catalyst layers much thicker than 10 μm, resulting, in turn, in limited catalyst utilization because of prohibitive charge and mass-transport limitations within catalyst layers of excessive thickness [10, 11, 13]. The significantly higher volume fraction of Pt in an optimized Pt black/ionomer composite versus that achievable in a Pt/C//ionomer composite, then becomes a significant advantageous feature of the black. The particle size of blacks is typically around 5 nm or larger; however, the metal volume density advantage overweighs the dispersion advantage of the supported catalyst when the target is a packing of higher catalyst surface area ($\geq$1 mg cm^{-2} of membrane surface), within a catalyst layer of limited thickness.

The electrochemical processes in the fuel cell take place at the interface between the dispersed anode and cathode electrocatalysts and hydrated ionomer electrolyte. In the hydrogen/air or reformate/air fuel cell, the processes at the anode and cathode, respectively, are as follows:

$$2H_2 = 4H^+ + 4e \qquad (2)$$

and

$$O_2 + 4e + 4H^+ = 2H_2O \qquad (3)$$

In the direct methanol polymer electrolyte cell (methanol/air fuel cell or DMFC), the anode process is as follows:

$$CH_3OH + H_2O = CO_2 + 6H^+ + 6e \qquad (4)$$

and the cathode process:

$$1.5O_2 + 6H^+ + 6e = 3H_2O \qquad (5)$$

In addition to the interfacial catalytic activity required to achieve and sustain a certain current density at a design cell voltage and temperature, the reactant mass transport has to be facilitated to maximize the overall rate of the fuel-cell electrode process, which is practically always under ''mixed control'' by rates of interfacial and mass-transport processes. The porous backing layer placed behind the catalyst layer (Fig. 1) enables effective gaseous reactant supply to the catalyst layer. Wet proofing by PTFE ensures that at least part of the pore volume in the cathode backing remains free of liquid water in an operating fuel cell, enabling high rate, gas-phase transport practically all the way to the catalyst surface. The scale of the porosity and the amount of PTFE added are important parameters that determine the effectiveness of the backing layer in fulfilling the combined tasks of gas and water transport, to and from the catalyst layer, respectively. Obviously, the backing layer (GDL) has to be made of material of high and stable electronic conductivity in the electrochemical environment of the fuel cell. Most of PEFC backing layers to date have consequently been based on wet-proofed porous carbon paper or cloth.

The outermost element of the unit cell (Figs 1 and 2) is the current collector plate, which typically contains a machined or molded gas flow field. These two functions of current collector and gas flow field may be fulfilled, in principle, by two separate components; however, in most of the cells and stacks built to date, the flow field is introduced as part of the current collector fabrication process. A current collector plate, with machined or molded flow field on each of its sides, becomes the

"bipolar plate" in the PEFC or DMFC stack (Fig. 2). The flow field has a wide range of possible geometries, for example, a single serpentine channel, parallel channel flow, and series-parallel combinations. This geometry is highly significant in fulfilling the requirements of uniform supply of gases across electrodes of large PEFC stacks of active areas of 250–1000 cm^2, together with effective water supply and/or water removal. Serving as an electronic current conduit and allowing fuel flow along one of its main surfaces and airflow along the other, the bipolar plate material has to exhibit high electronic conductivity and be impermeable to oxygen and fuel. At the same time, it has to be built as thin and light as possible, to achieve stack power density and specific power targets. Graphite, carbon/plastic composites, and metals like stainless steel and titanium, with or without surface coating or treatment, have been considered as bipolar plates/flow fields in PEFCs and DMFCs. The important advantage of the carbon-based materials is their good immunity to corrosion in the PEFC and DMFC electrochemical environments, whereas their cost, particularly of molded composite plates, can in most cases be brought down sufficiently with larger-scale manufacturing. The main drawbacks of carbon-based composites as bipolar plate materials are limited mechanical strength, which calls for a significant minimum plate thickness, and nonnegligible electronic resistance, particularly contact resistances that could arise as a result of surface depletion of the carbon component in a carbon/plastic composite. The latter problem can be avoided by using plates with a continuous graphite phase. One type of such bipolar plates, commercialized by Graftech (United States), uses natural graphite as raw material and allows molding plates with flow fields on both sides. Metal bipolar plates have some important advantages over carbon-based materials: better mechanical properties and higher electronic conductivity. Consequently, stacks based on metal hardware can achieve tighter packaging by using thinner bipolar plates and, therefore, achieve higher stack power densities. However, corrosion susceptibility of metal hardware in PEFCs and DMFCs, remains a challenge. Conflicting reports have described, on the one hand, successful implementation of metal hardware for PEFC and DMFC stacks and, on the other hand, testing showing insufficient immunity to corrosion. It seems that the choice of a specific metal alloy or surface-coated metal for PEFC hardware may be critical in achieving long-term, trouble-free performance. Reference 14 summarizes information on performance and durability of graphite and carbon composite bipolar plate materials, in addition to metal hardware including primarily stainless steel and surface-coated titanium and aluminum.

An MEA/bipolar plate pair is the building block repeat unit of the fuel-cell stack. The length of the stack, that is, the number of MEA/bipolar plate pairs stacked in series is determined by the source voltage required by the relevant load, with this number reaching, for example, 300 cells in transportation applications which require a 200-V power source. The surface area of each cell in the stack is then determined by the maximum power demand, for example, in such a 300 cell, 200-V stack, a cell active area of 250 cm^2 ($\frac{1}{4}$ ft^2) generates 50 kW of total power at current density of 1 A cm^{-2}, achievable at a PEFC single cell voltage of 0.66 V. The PEFC has been developed to date for a wide variety of applications, ranging from 50 to 200 kW stacks



Fig. 3 Examining a large-scale, framed MEA for a 75-kW PEFC stack (courtesy of UTC Fuel Cells).

for primary power sources in transportation all the way down to 1-W power sources for portable power in consumer electronics applications. These different applications obviously result in a wide variety of stack sizes and architectures.

An idea about stack sizes associated with potential transportation applications is given in Figs 3–5. Figure 3 shows an MEA mounted in a frame, to be used in a PEFC stack for transportation applications. Figure 4 shows the assembly of such a PEFC stack. It can be seen that hundreds of cells are stacked in series to provide the voltage requirement in an electric power train in a vehicle, typically over 250 V. Figure 5 shows a 75-kW PEFC developed for vehicular applications, in testing before installation in the vehicle. In (big) contrast, Fig. 6 shows the dimensions of a 1-W DMFC ''stack'' which, in this case, is actually not a stack in the ordinary sense, but rather a planar array of single cells connected in series electrically by current conduits external to the cells; a platform described as ''monopolar'' (as opposed to bipolar) series connection. The DMFC is ''air breathing'', that is, it is designed to be supplied by only a small flow of methanol fuel on one side, whereas air is supplied by natural convection and diffusion alone to the other side of this array of cells. It can be readily understood (see Fig. 4) that the design for air-breathing cathodes calls for a flat, monopolar array of cells to achieve effective air supply to all cell cathodes without resorting to an air mover of any type.



Fig. 4 Stacking the bipolar plates and MEAs for a 75-kW PEFC stack (courtesy of UTC Fuel Cells).

8.3.3 The Polymer Electrolyte and Direct Methanol Fuel Cells: Power Output and Energy-conversion Efficiency

The two primary merit parameters of a fuel-cell power source are the power output



Fig. 5 A 75-kW stack under testing (courtesy of UTC Fuel Cells).



Fig. 6 The molded flat array of air-breathing DMFCs, shown held between the fingers on the right, generates power in the range 0.5–1 W in a concept prototype of a DMFC system for portable consumer electronics applications. (MTI Micro Fuel Cells).

and the energy-conversion efficiency. The power output is usually gauged in terms of power density (kW L^{-1}) and/or specific power (kW kg^{-1}) of the stack, including the air, fuel, and water-management components, reflecting the ability to ''package'' some required power generating capacity into volume and/or weight restrictions set by some application. An example could be the packaging of a primary, 80-kW (109-HP) vehicular power source, under the hood of a passenger car, with the volume allowed for the fuel-cell stack limited to 60 L. This particular demand would require a stack of power density of 1.33 kW L^{-1}. PEFC-based power sources have become attractive technology for transportation applications owing mainly to demonstrated power densities of the order of 1 kW L^{-1}, similar to power densities of internal combustion engines. Fundamentally, this high performance of the PEFC is enabled by high interfacial activities at the catalyzed surfaces of the hydrated polymeric membrane and by the thin dimensions of the unit cell. The latter is enabled, in turn, by robust ionomeric membranes of thicknesses as

low as 25–50 μm and proper design and fabrication of thin-profile bipolar plates of sufficient mechanical strength. Beyond the quality of the catalytic electrodes and the key cell materials, the power density ultimately achieved in a PEFC or a DMFC stack depends further on quality stack engineering. Better engineering should, for example, maximize the active fraction of the total volume of a stack through reduction of volume and weight of manifolds and of structural components required to render mechanical strength and tight compression along the axis of the stack.

The combination of high intrinsic interfacial activity at the Pt/ionomer interface and effective catalyst dispersion also provides a key for high energy-conversion efficiency. The overall efficiency of a fuel cell converting fuel to electric power at some voltage, V_{cell}, across the load, is calculated as:

$$\eta_{\text{cell}} = \eta_{\text{fuel}} \times \eta_{\text{v}} \tag{6}$$

where η_{fuel} is the so-called fuel efficiency or fuel utilization, defined as the fraction of fuel flow into the cell converted to electric current, whereas the voltage efficiency η_{v} is defined as the cell voltage at the design point of operation divided by the thermodynamically allowed maximum voltage of the cell, V^0_{cell}

$$\eta_{\text{v}} = V_{\text{cell}}/V^0_{\text{cell}} \tag{7}$$

where V^0_{cell} is calculated as

$$V^0_{\text{cell}} = \Delta G^{\circ}/nF \tag{8}$$

The value of V^0_{cell} for a hydrogen/air fuel cell, 1.23 V at 25 °C and unity activities of hydrogen, oxygen, and water would be further corrected by the Nernst equation for specific values of p_{H_2}, p_{O_2}, a_{H_2O}, and a specific cell temperature. The corresponding Gibbs free energy change for the process in the methanol/oxygen cell, that is, methanol + oxygen = water + CO_2, corresponds to 1.21 V under standard conditions of 25 °C and unity activity of reactants and products.

The use of V^0_{cell} derived from the Gibbs free energy change (Eq. 8) in such calculations of the voltage efficiency in PEFCs is the accepted practice. However, it can be argued that the maximum theoretical conversion efficiency in a H_2/O_2 fuel cell should be calculated not on the basis of ΔG, but rather of ΔH, the enthalpy change in the process, as the latter reflects the total energy content of the fuel. The enthalpy change in the conversion: fuel + oxygen = water (+CO_2 in the case of methanol fuel), is calculated for temperatures significantly lower than 100 °C with consideration of water in *liquid* form as final product. The relevant value of ΔH° then corresponds, for the hydrogen/oxygen couple, to 1.48 V, rather than 1.23 V. The use of 1.23 V for V^0_{cell} in calculation of the voltage efficiency for a PEFC therefore requires some justification. One argument is that, in reality, under the active airflow along the cell cathode required to maintain cell performance, all the water product is likely to exhaust the cathode as vapor, rather than liquid, at a common cell temperature near 80 °C. An additional argument of different nature is that, in conversion efficiency calculations for the same fuel in combustion processes, the ''low heating value'' of the fuel is used consistently, that is, the heat released when water *vapor* is the final product. The ''low heating value'' of the fuel is practically equivalent to the Gibbs free energy change in the PEFC process and, therefore, the use of Eq. (8) to calculate V^0_{cell} amounts to

adaptation of the same "yardstick" for conversion efficiencies in the two types of fuel conversion processes, electrochemical and combustive.

As seen from Eqs (6) and (7), the overall efficiency of the fuel-cell scales linearly with the voltage across the cell under load and for hydrogen/air fuel cells, where η_{fuel} is very close to 100% and the overall cell efficiency is practically equal to η_{v}. Fuel-cell technology therefore ascribes high value to maximizing the cell voltage at any specified power demand. As would be expected, there is, however, a trade-off between the two main merit parameters of the fuel cell, conversion efficiency and power output. To evaluate the expected dependence of fuel-cell power on conversion efficiency, the cell "polarization curve" (i.e., the dependence of the cell voltage, V_{cell}, on the cell current, J) is the starting point. The latter is given, in general terms, by the following expression:

$$\begin{aligned} V_{\text{cell}} &= V_{\text{cath}}(J) - V_{\text{an}}(J) - J R_{\text{cell}}(J) \\ &= V^0_{\text{cell}} - \eta_{\text{cath}}(J) - \eta_{\text{an}}(J) - J R_{\text{cell}}(J) \end{aligned} \tag{9}$$

where $V_{\text{cath}}(J)$ and $V_{\text{an}}(J)$ describe electrode potentials at current density J of the cathode and the anode respectively versus a reversible hydrogen electrode (RHE), $\eta_{\text{cath}}(J)$ and $\eta_{\text{an}}(J)$ describe the overall voltage loss at current density J_{cell} at the cathode and anode respectively (calculated versus the thermodynamically expected electrode potential versus RHE at $J_{\text{cell}} = 0$), and R_{cell} is the high-frequency component of the resistance of the cell, determined by the sum of (1) the membrane resistance (which, under some circumstances, exhibits dependence on cell current J), (2) the cell plate/current collector bulk resistance, and (3) the various contact resistances originating from stacking and pressing together all cell components.

Following the general form of Eq. (9), a voltage–current density polarization curve for a hydrogen/air PEFC can be described in more specific terms by Eq. (10).

$$\begin{aligned} V_{\text{cell}} &= V_0 - \frac{RT}{F}\left(\frac{J_{\text{cell}}}{a_{\text{an}} J_{0,\text{an}}}\right) \\ &\quad - b_{\text{cath}} \log_{10}\left[\frac{\left(\dfrac{J_{\text{cell}}}{a_{\text{cath}} J_{0,\text{cath}}}\right)}{\left(1 - \dfrac{J_{\text{cell}}}{J_{\text{lim,cath}}}\right)}\right] - J R_{\text{ohm}} \end{aligned} \tag{10}$$

To highlight the major features of the PEFC polarization characteristics and their electrochemical origins, Eq. (10) is written assuming simple behavior of the fuel-cell anode and cathode behaving as ultrathin electrodes of total specific catalyst surface area (square centimeter of Pt per square centimeter of cross-sectional area) set by mass loading and dispersion of the catalyst. Additional losses in a PEFC electrode due to transport through the thickness dimension of the catalyst layer (see Sect. 8.3.7.2C) are not considered in Eq. (10) for the sake of simplicity. The equation can be seen to assume a constant "Tafel slope", b_{cath}, for the interfacial oxygen electrode process through the current range covered and a voltage loss at the hydrogen anode calculated from the relevant linear relationship between η_{an} and J_{cell}, which applies because the cell current densities (per square centimeter of Pt) are of the same order or smaller than the exchange current density of the anode process. The PEFC-limiting current is set by the air cathode limiting current, $J_{\text{lim,cath}}$ which is generally determined, in turn, by the rate of (gas phase) diffusion of oxygen from air through a porous cathode-backing layer of about 300-μm thickness, with the pores filled

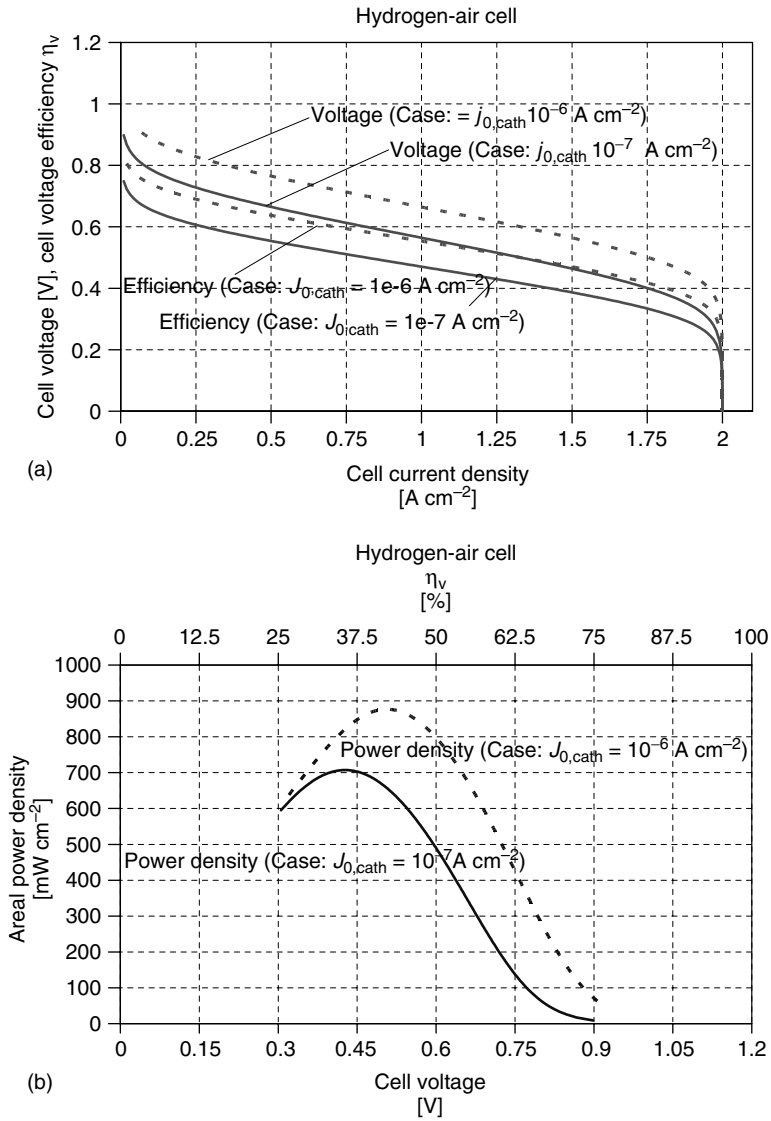


Fig. 7 (a) Simulated voltage/current density and efficiency/current density relationships for a hydrogen/air PEFC based on Eq. (10), assuming the following parameters: $J_{0,an} = 0.05$ A cm^{-2} Pt, $J_{0,cath} = 10^{-7}$A cm^{-2} Pt, or 10^{-6}A cm^{-2} Pt, $a_{an} = a_{cath} = 100$ cm^2 Pt/cm^2 geo, $b_{cath} = 100$ mV/decade of current density, $J_{lim,cath} = 2.0$ A cm^{-2} and $R_{ohm} = 0.1\ \Omega$ cm^2. (b) Areal power density plotted against cell voltage and against cell conversion efficiency, for PEFCs of current voltage characteristics shown in Fig. 7(a).

by nitrogen and water vapor [13]. Specific values of the parameters in Eq. (10) used to generate Fig. 7(a) below were $J_{0,an} = 0.05$ A cm^{-2} Pt, $J_{0,cath} = 10^{-7}$ A cm^{-2} Pt, or 10^{-6} A cm^{-2} Pt $a_{an} = a_{cath} = 100$ cm^2 Pt/cm^2 geo, $b_{cath} = 100$ mV/decade of current density, $J_{lim,cath} = 2$ A cm^{-2} and $R_{ohm} = 0.1\ \Omega$ cm^2. These electrode parameters are typical for electrodes of low Pt loadings of 0.1–0.2 mg Pt/cm^2 of electrode area, operating near 80 °C, and the specific resistance is typical of the poly(PFSA) membranes used in PEFCs. Because the cathode voltage loss is, by far, the most dominant, two levels of intrinsic catalytic activity are considered in the polarization curve based on Eq. (10): $J_{0,cath} = 10^{-7}$ A cm^{-2}Pt and 10^{-6} A cm^{-2}Pt, highlighting what such improvement by an order of magnitude in cathode catalytic activity could do to enhance the conversion efficiency/power density combination available from a PEFC stack. It is obvious that the same enhancement can be achieved with higher active catalyst surface area (higher a_{cath}), although the latter approach will incur higher cost. As can be seen from actual, measured PEFC polarization curves for operation on neat hydrogen and oxygen, and neat hydrogen and air, shown in Fig. 8, the calculated polarization curves in Fig. 7(a) reflect quite well the polarization characteristics of hydrogen/air PEFCs, employing thinner, state-of-the-art poly(PFSA) membranes. The curves shown in Fig. 8 exhibit, however, stronger voltage losses at the highest cell currents that are caused by transport limitations within the cathode catalyst layer, not considered in the simplified calculation (Eq. 10), which assumes ultrathin electrodes.

As the PEFC power output, per unit MEA surface area is given, at some cell voltage, by $P_{out} = V_{cell} \times J_{cell}$ and, since the increase in J_{cell} with drop in V_{cell} is superlinear through most of the useful cell voltage range (Fig. 7(a)), P_{out} increases with the rise in cell current/decrease in cell voltage, until further rise in cell current brings about the much sharper voltage drop caused by cathode mass-transport limitations (Fig. 7(a)). The consequence is an inverse relationship between cell power output, P_{out}, and cell voltage, V_{cell}, and, in cases where the fuel efficiency η_{fuel} is practically independent of cell current, the same inverse relationship would be observed between the power output, P_{out}, and the cell conversion efficiency η_{cell} (Eq. 7). Calculated P_{out} versus V_{cell} and versus η_{cell} for the polarization curves in Fig. 7(a) are given in Fig. 7(b). The case $\eta_{cell} = \eta_v$ (Fig. 7(b)) applies to the hydrogen/air PEFC; however, this is *not* the case for the methanol/air cell, where η_{fuel} can vary significantly with cell current under some given rate of fuel delivery to the anode. The dependence of P_{out} on η_{cell} in the DMFC is therefore quite different than that shown in Fig. 7(b) and is discussed in detail in the section 8.3.8.2 as evident from actual, measured cell performance (Fig. 8), the calculated power efficiency plots in Figure 7(b) reveal the typical range of areal power output/conversion efficiency combinations, available from a hydrogen/air PEFC. At peak power of 0.7 W cm^{-2} (0.7 kW ft^{-2}) of electrode cross-sectional area, the conversion efficiency is close to 50%, whereas at 25% of this peak power the conversion efficiency can be as high as 70%.

Conversion from areal power output of an MEA, P_{out}, expressed typically as W cm^{-2}, to stack power density P^*_{stack}, expressed typically in units of watts per cm^3 or kilowatts per liter, is determined by two factors: the width of a single unit cell in the stack ("cell pitch"), δ_{cell}, and the fraction

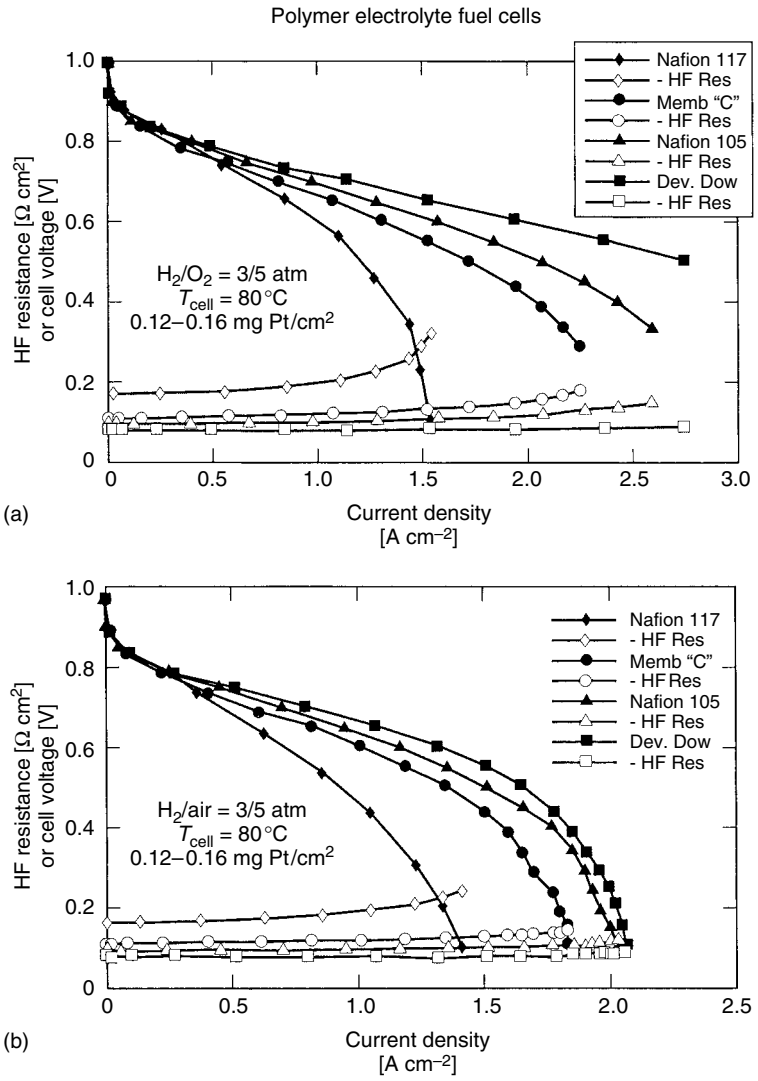


Fig. 8 Hydrogen/oxygen and hydrogen/air PEFC performance on low Pt catalyst loadings, as obtained at Los Alamos National Laboratory in the early 1990s using catalyst ''inks'' applied to the membrane surface. Cell characteristics are shown for several poly(PFSA) membranes of different thicknesses, with catalyst loadings in the range of 0.12–0.16 mg Pt/cm^2 applied to each side of the membrane. Notice: (a) the current limit of 1.5 A cm^{-2} in the case of the thicker membrane is clearly associated with strong increase in membrane resistance with cell current in the case of the Nafion 117 membrane, which is lifted for thinner membranes (see Sect. 8.3.4.4.3); (b) the current limit of 2 A cm^{-2} for operation on air (but not on oxygen) with thinner membranes is caused by oxygen transport limitations as a result of dilution in the inert mixture of nitrogen and water vapor.

of active volume out of the total volume of the stack, f_{active}. This fraction enhanced by minimizing nonactive components of the stack, including end plates, fuel and air feed plenums, and extra components devoted to stack humidification, for example, "enthalpy exchangers". P^*_{stack} is given, in general, as

$$P^*_{stack} = P_{out}[1./\delta_{cell}]f_{active} \qquad (11)$$

As an example, at 0.6 W cm^{-2} of MEA area, unit cell width of 3 mm, and active volume fraction in the stack of 60%, the volume power density of such stack is calculated as 1.2 W cm^{-3} or 1.2 kW L^{-1}.

Considering the very low exchange current density of the ORR process in aqueous acid electrolytes, the most significant single achievement leading to such high-performance hydrogen/air PEFCs, is the PEFC air electrode technology. Resolving the cathode voltage loss and tying it to the interfacial activity helps clarify how such high air electrode activity is achieved. This requires determination of the cathode potential, $V_{cath}(J)$, at cell current J, which requires, in turn, correction of the measured cell voltage, $V_{cell}(J)$, for ohmic loss. The ohmic component of the cell impedance, R_{cell}, is measured as the high-frequency component of the impedance (1–10 kHz) with the precise value of the frequency best selected by a phase angle of zero degrees (i.e., no imaginary component). An "iR-corrected cell voltage" can then be calculated at each measured current density, from the measured V_{cell} and R_{cell}, according to

$$V_{corr}(J) = V_{cell}(J) + J\,R_{cell}(J) \qquad (12)$$

In a PEFC with a pure hydrogen anode feed, $V_{cath}(J)$ is practically identical to $V_{corr}(J)$, as the anode overpotential η_{an} is, at the most, 10 mV at a cell current density of 1 A cm^{-2}. With V_{cell} near 0.60 V at cell current density of 1 A cm^{-2} – a typical PEFC peak-power point at $T_{cell} = 80\,^{\circ}$C- and with R_{cell} of typically 0.1 Ω cm^2, V_{cath} would be given to good approximation as $V_{cath} = 0.6\text{ V} + J\,R_{cell} = 0.70$ V versus RHE. At $V_{cath} = 0.70$ V, η_{cath} is over 0.5 V and the ORR current density per square centimeter Pt is expected to be 4–5 orders of magnitude higher than the exchange current density (see Eq. 10), that is, $J_{cell} =$ 10 mA cm^{-2}Pt. Consequently, a cell current of 1 A cm^{-2} of MEA cross-sectional area will be obtained at $V_{cath} = 0.70$ V with 100 cm^2 of catalyst area per square centimeter of MEA cross-sectional area. This degree of "surface area amplification" is achievable at 0.1–0.2 mg Pt/cm^2 of MEA, assuming a Pt particle size of 2 nm and a high catalyst surface utilization. In summary, a low ohmic loss and good technology of catalyst dispersion and utilization enables to reach in a PEFC an areal power density of 0.6 W cm^{-2} together with a conversion efficiency of 40–50%, in spite of the intrinsically low activity of the air electrode ($J_{0,cath} =$ 10^{-7}–10^{-6} A cm^{-2} Pt). The cathode voltage loss, η_{cath}, of over 0.5 V, required to reach the maximum power point in the PEFC, corresponds, however, to a loss of over 40 percentage points ($0.5_5/1.23$) in voltage efficiency, that is, in conversion efficiency. Another smaller loss of eight percentage points (0.10/1.23) is incurred by the PEFC's ohmic resistance, associated with the membrane electrolyte itself and with inner cell contact resistances.

In the case of the DMFC, a typical design point of $V_{cell} = 0.45$ V, in a state-of-the-art cell, would allow a current density of 0.2 A cm^{-2} at cell temperature of 80 °C. Under these conditions, the anode potential would typically be 0.4 V versus

RHE, the cathode potential 0.85 V versus RHE, and the iR loss (at 0.2 A cm^{-2}) would typically be 0.03–0.05 V. In the case of the DMFC, the anode and cathode contributions to cell voltage loss, and hence to efficiency loss, are comparable, each causing, at an areal power density $P_{out} = 0.1$ W cm^{-2}, an voltage efficiency loss of around 30%. With the additional iR loss of 4%, the resulting overall voltage efficiency is $\eta_v = 36\%$. Unlike the case of the PEFC, the fuel efficiency, η_{fuel}, in the case of the DMFC could be quite lower than 100%, and, to calculate the overall efficiency, this factor also has to be carefully considered (Eq. 6). At $\eta_{fuel} = 90\%$, which is achievable by proper design of cell and operating conditions as discussed in Sect. 8.3.8.2, the overall DMFC stack conversion efficiency at peak power would be $\eta_{cell} = 32\%$. Operation of the DMFC at half the maximum power output (0.05 W cm^{-2}) could raise the cell voltage to 0.50 V and, consequently, raise η_v to 42%, and, with $\eta_{fuel} = 90\%$, the DMFC conversion efficiency will be $\eta_{cell} = 38\%$.

With areal power outputs only 20–30% that of a PEFC and an energy-conversion efficiency of 30% near peak power versus 50% in the case of the PEFC, the DMFC remains of great interest because of the attractive properties of methanol fuel, a liquid of high energy density under ambient conditions, and because the DMFC enables direct conversion of this liquid carbonaceous fuel to electric power. Particularly in portable applications, these features help minimize the overall dimensions of the power system (fuel + fuel cell + auxiliaries) and achieve high system energy density.

Looking at both the PEFC and DMFC state of the art, one can readily recognize the "empty half of the glass", side by side with the "full half of the glass". While exceeding at such low temperature the conversion efficiencies of internal combustion engines at similar volume power densities the voltage losses in the PEFC and DMFC at useful power densities remain substantial. As explained earlier in this section, these losses originate from limited interfacial kinetics and limited transport rates, as well as from membrane electrolyte properties. The fundamental nature of cell components and materials and cell losses associated with each are discussed in more detail in the following sections.

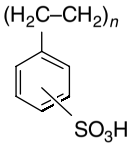
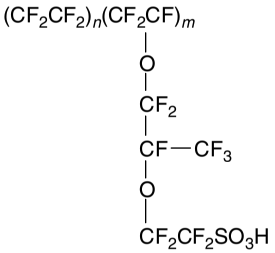
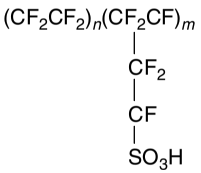
8.3.4 The Ionomeric Membrane

8.3.4.1 Chemical Structure and Chemical Stability of Proton-conducting Polymeric Membranes

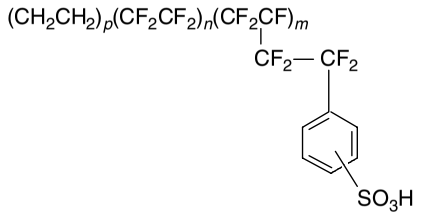
PEFCs are defined as a special class of fuel cells owing to the polymeric nature of the electrolyte. This electrolyte confers on the PEFC unique properties of a pseudo-solid-state device (only deionized water needs to be managed), operating at relatively low temperatures. The properties of the polymeric membrane are, therefore, the key for achieving high cell performance and long-term stability. Chemical structures of a variety of membrane materials that have been used in PEFCs, or developed and tested for PEFCs, are shown in Table 1.

Initially, poly(styrenesulfonic acid) (PSSA) and sulfonated phenol-formaldehyde membranes were used for PEFCs, but the useful life of these materials was limited because of significant degradation under fuel-cell operating conditions. A critical breakthrough was achieved with the introduction of Nafion®, a perfluorinated polymer with side chains

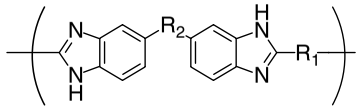
Tab. 1 Chemical structures of Ionomeric Membranes Employed in (or considered for) PEFCs and DMFCs

Membrane	*Chemical structure*
PSSA	
Nafion	
Poly(PFSA) of shorter chain as in the membrane, first presented by Dow Chemical. Similar membrane is now fabricated elsewhere	

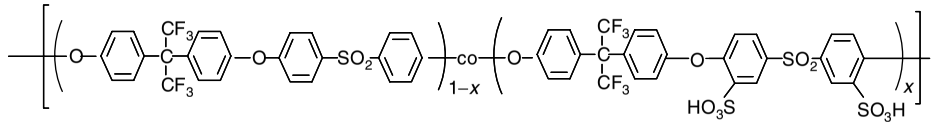
α, β, β-Trifluorostyrene grafted onto poly(tetrafluoroethylene-ethylene) with postsulfonation



Polybenzimidazole – the polymer used as membrane in the fuel cell is intermixed with neat phosphoric acid



Sulfonated polyaromatics (partly fluorinated in this specific case), tested for possible operation at $T_{cell} > 100\,^{\circ}C$ and/or lower methanol permeability



terminating in sulfonic acid moieties. This specific material and its close poly(PFSA) relatives are currently the state of the art in membranes for PEFCs, satisfying, to a significant degree, the requirements for effective use in fuel cells, albeit with some remaining outstanding longevity issues. The poly(PFSA) membranes combine well the important requirements of high protonic conductivity, relatively high chemical stability under relevant operating conditions, and low gas permeabilities. Continuous tests of PEFCs based on such membranes under constant load for 10^4 hours and more with limited performance loss have been documented [13]. In spite of extensive work on alternative proton-conducting membranes for the PEFC, poly(PFSA) membranes have maintained to date a position of advantage versus any alternative, established due to the combined properties of high protonic conductivity and superior chemical and mechanical robustness. A central source of PFSA membranes is DuPont (United States), where these membranes were invented in the 1960s, used first in the 1960s for the fuel cell in the Gemini space program and made first into a commercial product (in Na^+ form) for the chlor-alkali industry. Other sources of PFSA membranes for fuel cells have been Asahi Glass (Japan), Asahi Chemical (Japan), and 3M (United States). Microcomposite membranes for fuel cells based on PFSA as the proton-conducting material have been developed at and supplied by W.L. Gore (United States). The latter membrane has a backbone made of expanded PTFE, providing better mechanical strength particularly for thinner membranes (<50 μm) and limiting membrane dimensional changes following repeated dry–wet cycles. A PFSA component of the Gore membrane, apparently filling the porous network in the reinforcing, expanded PTFE backbone, provides protonic conductivity of the same order as that obtained from homogeneous PFSA membranes.

The three-dimensional structure of some of these poly(PFSA)membrane materials has been extensively studied and reviewed [14–16]. Eisenberg and coworkers [14] showed first that ion aggregates form in poly(PFSAs). The energetics of cluster formation have been discussed in Ref. 17. The existence of these ionic nanoclusters is generally agreed upon. Details of their size and distribution and the effective porosity and tortuosity of the resulting aqueous/ionic nanoscale network that forms from them within the membrane have been recently revealed by small angle X-ray spectroscopy (SAXS), particularly the tapping mode (TM) atomic force microscope (AFM) work [18]. The latter technique demonstrated ionic domains in Nafion membranes exposed to normal ambient humidity, which exhibit high degree of uniformity in spacing, with an approximate size of 4–10 nm in diameter.

Images of water-swollen membranes presented clumps of ionic regions that increased to 7–15 nm in size while appearing to have coalesced into channel-like shapes. Recent work by Weber and Newman [19] suggested that contact with liquid water is a prerequisite for the significant widening of the interaggregate channels, thereby explaining the significant further rise in water and proton mobility in poly(PFSA) membranes on transition from saturated water vapor to liquid water in contact with the membrane.

In spite of their superior characteristics for PEFCs, PFSA membranes do have some weak points. Firstly, their relatively high cost (500–1000 $\$/m^2$) has inspired

continued work on lower cost alternatives. It is unclear how much the cost of poly(PFSA) membranes could drop if the demand for them increases, but it has been argued that a 2 order-of-magnitude increase in the market for such membranes – their market is now limited to chlor-alkali industry application and to development, demonstration, and testing of PEFC power system prototypes – will bring about an order-of-magnitude decrease in their price. For transportation, considered the largest potential market for PEFCs in the longer run, the target power system cost is under 50 $/kW and that would require the lowering of membrane cost to well under 50 $/m^2. A drawback of a different nature is the limited temperature range over which these membranes can be effectively used. The upper temperature limit is usually considered around 100 °C because of the low membrane water content above that temperature under low internal over-pressure. Above 100 °C the water content in a poly[PFSA] membrane required to maintain high proton conductivity, can be achieved only under significant water vapor pressure elevated total pressure requiring, in turn, maintained in the fuel-cell stack. The low limit of the useful temperature window is set by the strong fall of the protonic conductivity in a hydrated PFSA membrane under the freezing point of water. The inner PEFC temperature window, enabling high performance, is consequently considered 0–100 °C (preferably 10–80 °C) and, in the design of a PEFC stack subsystem, provision has to be made to raise the cell temperature above the freezing point of water when the ambient temperature is lower than that, and proper shedding of heat generated by the cell process is required to limit the rise of the cell temperature to a maximum near 80 °C.

For optimal use in fuel cells, in addition to the primary requirement of high protonic conductivity, the membranes employed in PEFCs must be chemically and mechanically stable in the fuel-cell environment. The chemical stability requirement has been an important issue since the initial development of membranes for PEFCs. Degradation of PSSA membranes used initially for PEFCs in the early 1960s is severe even under conditions significantly milder than those encountered in an operating fuel cell [20, 21]. Consequently, the introduction of the poly(PFSA) membranes, where C–H bonds are fully replaced by C–F bonds in the polymer structure, enabled an important step-up in membrane stability in the operating PEFC. Subsequent attempts to develop alternative membranes, not based on perfluorinated polymers, have included, for example, ionomers with poly(aromatic) backbones and fluorinated, sulfonic acid–terminated sidechains. The latter membranes, as a rule, have lower protonic conductivity at a given degree of sulfonation, or given equivalent weight (EW), because of smaller tendency for the nanoscale, aqueous ionic phase separation, which occurs most readily in ionomers with highly hydrophobic, perfluorocarbon backbone. Otherwise, such ''hydrocarbon membranes'' show, as a rule, lower chemical stability versus that exhibited by ''perfluorinated membranes''.

In spite of the documented, relatively high chemical stability of poly(PFSA) membranes in the fuel-cell environment, recent extensive work looking into the origins of performance loss observed in PEFCs has revealed important mechanisms of degradation that apply to perfluorinated membranes (while being further amplified in nonperfluorinated membranes). An important mechanism of membrane

deterioration identified, is attack by OH radicals that tend to form by recombination of oxygen and hydrogen, taking place typically by oxygen crossing over through the membrane to the anode and reacting there with hydrogen, or vice versa. Following such initial attack by $OH^{\bullet}$ radicals, further poly(PFSA) membrane attack can occur by hydrogen peroxide, a product of the PEFC cathode in the presence of various adsorbed impurities. This mechanism was, in fact, revealed during early stages in the development of PEFCs with poly(PFSA) membranes [20], but was not considered, until recently, too significant in new generations of commercial poly(PFSA) membranes and when hydrogen/air, rather than hydrogen/oxygen cells, are considered. Improved poly(PFSA) membrane synthesis/preparation is being sought at this point, which would further lower this vulnerability to $OH^{\bullet}$ radical and hydrogen peroxide attack. Also, cerium and manganese ion additives were very recently reported to have significant protective capacity against OH radical attack [PFSA] Fuel Cell membranes.

A significant amount of recent work on alternative PEFC membranes has targeted operation at higher temperatures, in the range 120–200 °C, and resulted in several new developments, the most significant of which is probably the phosphoric acid (PA)-doped poly(benzimmidazole) (PBI) [21]. In the PA-doped PBI, significant proton conductivity is maintained at temperatures in the range 120–160 °C at low water activities, apparently owing to proton hopping sustained by the PA ''dopant'' (neat PA is known to allow effective proton hopping, utilizing oxygen atoms on adjacent PA molecules as hopping sites). Low levels of water added to the PA-doped PBI facilitate dissociation, but the water content in question is low and does not require significant steam pressure, thereby enabling PEFC fuel cell operation at such higher temperatures at practically ambient pressure. Blending PA with PBI is assisted by acid–base interactions between the acid and the base functions of the PBI building block (see Table 1). Nevertheless, the composite, PA-doped PBI and the poly(PFSA) membranes belong to different categories of polymer electrolytes. The latter uses only distilled water, whereas the former relies on some liquid electrolyte blended/bonded with a polymeric matrix. If the composite membrane blend is stable under some relevant operation conditions, the ''dopant'' electrolyte can provide some important advantages, rendered by a proton-conducting fuel cell operating at over 120 °C under low pressure. These include much better impurity tolerance than in operation at or under 80 °C and more effective heat rejection. Such electrolytes would therefore be considered when significant system simplification is to be achieved in operation on some processed carbonaceous fuel. In such cases the much better impurity tolerance of the fuel cell allows to eliminate fuel processing stages required to strongly lower the carbon monoxide level in the anode feed. This advantage at the system level could be sufficient to compensate for some drawbacks of the PA-doped PBI electrolyte, including the well-documented lowering of the electrocatalytic activity of a Pt catalyst when in contact with PA electrolyte. This lower catalytic activity results in hydrogen/air cells with PA-PBI membrabe electrolyte exhibiting, at some given current density, cell voltages that are typically 200–300 mV lower versus hydrogen/air PEFCs employing poly(PFSA) membranes.

8.3.4.2 The Ionomeric Membrane/Water System

The protonic conductivity of an ionomeric membrane is strongly dependent on membrane water content. The membrane water content at a certain coordinate within an operating PEFC is determined, in turn, by two characteristics: (1) the dependence of membrane water content on the local water activity in contact with the membrane surface at the relevant temperature and (2) water boundary conditions and water fluxes generated in the operating cell by gradients in chemical, electrical, and/or mechanical energy, which together determine a steady-state water activity level at any specific coordinate within the cell under given cell operation conditions.

8.3.4.2.1 Water Uptake by Ionomeric Membranes Depending on the overall local water content and temperature, the ionomeric membrane can be exposed within an operating PEFC to water in either vapor form only, or in mixed, liquid and vapor form. There are significant differences between these two cases regarding the equilibrium water levels in the membrane as well as the dynamics of water uptake by the membrane.

The maximum level of water uptake by the membrane on immersion of a poly(PFSA) membrane in liquid water ranges between 20 and 50% by weight, depending primarily on the EW of the membrane. This corresponds to 20–25 water molecules per SO_3^- group (or per proton) in the membrane. This high number of water molecules taken up per proton suggests that the local environment of the proton in a fully hydrated poly(PFSA) membrane is not very different than in aqueous acid solutions, particularly when considering the tendency of these membranes to phase separate and form ion/water nanoclusters. Clearly, the ability of these membranes to imbibe that much water per ionic group while maintaining mechanical integrity and low gas permeability, is an important key to their successful use in PEFCs of high power output. The dependence of water uptake from the liquid on membrane thermal pretreatment was mentioned in the earliest descriptions of Nafion® properties. As pointed out early on in [22], Nafion membranes take up dramatically more water (up to 100% by weight) from liquid water at temperatures above 100 °C and, in contrast, membranes dried out completely at elevated temperatures imbibe significantly less water on reimmersion in liquid water at lower temperature (as low as 10% by weight). The origin of these phenomena lies in the combined effect of temperature and hydration on polymer structure. Some specific data on the different amount of water taken up by poly(PFSA) membranes immersed in liquid water with no pretreatment and with thermal pretreatment is given in [22, 23]. As seen in Table 2, the relative water uptakes, expressed in weight percent, indicate the strong effect of predrying at 105 °C on subsequent water uptake form the liquid. The table also shows a substantial increase in water uptake with decrease in membrane EW (= increase in volume density of charge carriers). Poly(PFSA) membranes of lower EW are seen to take up more water per unit membrane weight and, consequently, should exhibit higher protonic conductivity. Further lowering of EW under $EW = 800$ could, however, jeopardize the mechanical integrity of the membrane.

While high membrane water contents can be established in contact with liquid water, water vapor can be the principal mode of hydration of the PEFC membrane under some, not uncommon operation conditions. The shape of isotherms of

Tab. 2 Water uptake by various membranes from liquid water, expressed as λ (moles H_2O/mole of SO_3H) and as percent of dry weight

	(1) No thermal treatment	*(2) Dried at 105 °C*[a]		
	Rehydration temperature	*Rehydration temperature*		
Membrane type	*27 °C<T<94 °C*	*27 °C*	*65 °C*	*80 °C*
Nafion™ 117[b]	21 (34%)	12 (20%)	14 (23%)	16 (26%)
Poly(PFSA), $EW = 900$	21 (42%)	11 (22%)	15 (30%)	15 (30%)
Poly(PFSA), $EW = 800$	25 (56%)	16 (36%)	23 (52%)	25 (56%)

[a] Membrane dehydrated completely at this temperature (incompletely dried membranes behave as in the ''No thermal treatment'' column).
[b] $EW = 1100$.

water vapor sorption by a poly(PFSA) membrane, determined by isopiestic measurements and shown in Fig. 9(a) and (b), is generic for ion-exchange polymers. With increasing P_{H_2O}, water is sorbed by the membrane in two distinguishable steps evidenced by the sorption isotherm: (1) a low vapor activity region, $a_{H_2O} = 0.15 - 0.75$, characterized by a relatively small increase in water content with water vapor activity and by enthalpy of sorption of about 12.5 kcal mol^{-1} (greater than the enthalpy of water liquefaction) and (2) a high vapor activity region, $a_{H_2O} = 0.75 - 1.0$, characterized by steeper increase in water content with water activity and by sorption enthalpies as low as 5 kcal mol^{-1} at $a_{H_2O} = 1.0$ (significantly lower than the enthalpy of water liquefaction). Region (1) corresponds to uptake of water of solvation by the ions in the membrane, while region (2) corresponds to water that fills (submicro) pores and swells the polymer. The lower enthalpy for the second step is an indication of a weaker water–ion interaction and the endothermic deformation of the polymer matrix as the polymer is swelling. The further water uptake and associated swelling of the membrane at higher water vapor activities must therefore be driven significantly by increase in entropy. From Fig. 9, it can be seen that differences in water uptake between different PFSA membranes, expressed in terms of percent of dry weight taken up, are almost completely explained by differences in EW. When water uptake is expressed in terms of the ratio λ = number of water molecules/sulfonate group, the uptake curves look essentially identical for the different PFSA membranes (Fig. 9(b)).

Comparison of water uptake by PFSA membranes from the liquid and from the vapor phase reveals an interesting apparent paradox: the water content of the membrane in equilibrium with saturated water vapor ($a_w = 1$) is lower than the water content of a similarly prepared membrane in contact with liquid water ($a_w = 1$). Under isopiestic conditions (Fig. 9(b)), 14 waters per sulfonate group are sorbed from vapor at unit activity (saturated vapor), whereas 22 are sorbed from liquid water at the same temperature. The phenomenon was first reported in 1903 by Schroeder [24] and is named after him as *Schroeder's paradox*. The explanation suggested for the lower water uptake from the saturated vapor is that sorption of water from the vapor

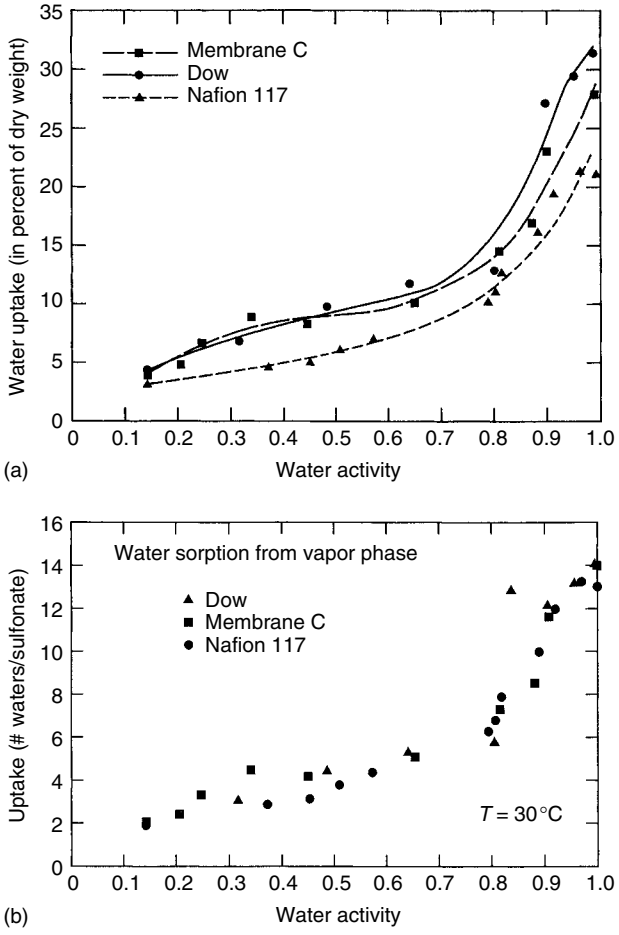


Fig. 9 Measured water uptake by three PFSA membranes from the vapor phase at 30 °C as a function of water vapor activity: (a) in terms of weight percent; (b) as the ratio of number of water molecules/number of sulfonic acid groups [111].

beyond the capacity of the aqueous, ionic clusters in the membrane must involve condensation of water vapor onto strongly hydrophobic (Teflon-like) pore surfaces within the polymer – an improbable process – whereas water penetration into such hydrophobic pores from liquid water in contact with the membrane can be driven by a gradient in liquid pressure and occur to the degree allowed by the deformability of the swelling membrane. This general picture has been articulated most recently in Ref. 19. A manifestation of barriers to water uptake by poly(PFSA) membranes from the vapor is provided by data showing a smaller water uptake, of only $\lambda = 10$ at higher membrane/water equilibration temperatures of 80 °C (corresponding to a typical operating temperature of a PEFC), to be compared with the higher water level

of $\lambda = 14$ seen at equilibration temperatures near 30 °C [25]. As might be expected, the probability of water condensation onto hydrophobic pore surfaces drops further at elevated temperatures.

8.3.4.2.2 **The State of Water in the Ionomeric Membrane** A variety of techniques have been applied to study the state of water in ionomeric membranes. The central goal has been to understand the consequences of the local structure and energetics of water in the membrane on water mobility which would have, in turn, a significant effect on proton mobility and thereby on the protonic conductivity of the membrane. Earlier differential thermal analysis studies [24] determined a ''nonfreezing'' fraction of water in ionomeric membrane materials and suggested that each of the sufonated polymers studied (polystyrene sulfonic acid (PSSA, polyethylene sulfonic acid (PESA), and Nafion, $EW = 1100$) has ''nonfreezing'' water at a level near $6H_2O/SO_3H$. An enthalpy of fusion of 5.65 kJ mol^{-1} was measured for water in those membranes beyond the ''nonfreezing'' level, similar to that of free liquid water -6.01 kJ mol^{-1}. However, below roughly six water molecules per sulfonate group, no enthalpy change was observed down to -40 °C. This measurement indicates that the ''first'' molecules of water taken up when a dry ionomeric membrane is exposed to water vapor exhibit characteristics that are quite different from those of free liquid water. This is readily understood by the strong bonding to ionic sites in the ionomer, expected for the ''first'' water molecules introduced into a dry membrane. Differentiation between water at those low hydration levels of up to $6H_2O/SO_3H$ and the further water taken up beyond that level, reaching $>20H_2O/SO_3H$ in a swollen poly(PFSA) membrane in contact with liquid water, has been at the center of studies of the poly(PFSA)/water system. Another clear thermodynamic differentiation is that membrane water at the lowest hydration levels is associated with enthalpy of water uptake larger than the heat of water condensation, whereas for water added into a membrane of high water content, the enthalpy of vapor uptake is significantly lower than the enthalpy of water condensation. This clarifies further that, at low water levels in the membrane, water molecules are strongly bonded to ions, whereas with further increase in water level, the water becomes more and more bulklike and the process of further hydration is driven by entropy rise rather than by enthalpy drop. The consequence for PEFC technology is that a strong difference in water (and hence proton) mobility is to be expected with membrane water content, with a ''demarcation line'' in water properties between the first five to six waters per sulfonate group and higher water content. A recent molecular dynamics work by Paddison et al. [26] shows that another factor strongly affecting protonic mobility, namely the tendency of the sulfonic acid groups to dissociate according to $-SO_3H + H_2O \rightarrow -SO_3^-$ (H_3O^+), also dramatically increases as water levels reach $6H_2O/SO_3H$. Consequently, the general picture of the aqueous subphase in the ionomeric membrane, the higher hydration levels targeted for high-performance PEFCs, is that of concentrated, well-dissociated sulfonic acid ($-SO_3^-$ (H_3O^+)) in aqueous medium of properties close to that of bulk water.

The water/membrane system behavior at temperatures below the freezing point of water is, for obvious reasons, very significant technologically. Some recent reports on self-start of PEFC stacks from

temperatures well below freezing are of, therefore, special interest. In this regard, it should be clarified that the absence of an observed enthalpy change around the water freezing point for a membrane with up to $6H_2O/SO_3H$ cannot be interpreted to mean that such ''nonfreezing form of water'' in the membrane can prevent strong loss of protonic conductivity as the stack temperature drops under 0 °C. As explained in the preceding text, these ''nonfreezing'' water molecules are strongly associated with ions and their mobility is, consequently, most limited. At temperatures well below 0 °C, membranes of higher water content would have such ion-bound molecules surrounded by a water phase akin to solid ice, suggesting, in turn, a strong drop in overall water mobility in the membrane under the freezing point. However, recent reports have revealed that a deeply frozen PEFC stack (e.g., idled for hours at an external temperature of −30 °C) is capable of self-heating all the way up to its typical operation temperature by passing current through the deeply frozen stack of cells. The mechanism of protonic conductivity in an ionomer under conditions of such low water translation mobility could, in fact, be different than that which applies under ordinary PEFC operation conditions. However, irrespective of a still-awaited full explanation of the proton conduction mechanism in such frozen stacks, the significance of such a demonstration of ''stack bootstrapping from −30 °C'' in establishing the viability of PEFC technology for transportation applications, is quite obvious.

8.3.4.3 Proton Conduction in Ionomeric Membranes

The most important physical property of ionomeric membranes employed in PEFCs is the high protonic conductivity they provide under proper humidification conditions. As in any conducting medium with a single-charge carrier, the magnitude of the specific conductivity is determined by the product of charge carrier density and charge carrier mobility as follows:

$$\sigma = FC\mu, \tag{13}$$

where σ is the specific conductivity ($S\,cm^{-1}$), F is Faraday's constant ($C\,mol^{-1}$), C is the concentration of the charge carriers ($mol\,cm^{-3}$), and μ is the charge mobility ($cm^2\,V^{-1}s^{-1}$). It is the proton mobility factor in Eq. (13) that is strongly affected by the membrane water content and the state of water in the membrane at the relevant water content. These two factors define together the mobility of water in the membrane as discussed earlier and, consequently, also define the mobility of the proton, as the latter moves through the membrane primarily by a ''vehicular mechanism'' with the water molecule serving as the ''vehicle''. The charge (proton) carrier density in an ionomeric membrane of $EW = 1100$ is similar to that in a 1.5 M aqueous sulfuric acid solution and, quite remarkably, the proton mobility in a fully hydrated poly [PFSA] ionomeric membrane is only 1 order of magnitude lower than the proton mobility in an aqueous solution. As a result, the specific conductivity of a fully hydrated (liquid immersed) poly(PFSA) membrane is 0.06–0.08 $S\,cm^{-1}$ at room temperature and 0.15–0.16 $S\,cm^{-1}$ at a typical elevated cell temperature of 80 °C. These levels of specific ionic conductivity are at least 2 orders of magnitude higher than specific ionic conductivities achieved for solvent-free, ionically conducting polymers at similar temperatures, thereby providing the basis for the high power densities achievable in PEFCs.

The conductivity of Nafion 117 decreases roughly linearly with decreasing water content, as shown in Fig. 10. The value of the conductivity measured at $\lambda = 14$ (contact with saturated vapor) at 30 °C is seen to reach 0.06 ± 0.01 S cm^{-1}. For the other two poly(PFSA) membranes of lower EW, the conductivity decreases roughly linearly with decreasing water content until an apparent threshold is reached, at which point the conductivity drops substantially (Fig. 10). Above the water content where the steep drop occurs, the conductivity of both membranes is much higher than that of Nafion 117. This is expected owing to the larger concentrations of charge carriers (= lower EW) relative to Nafion 117 (see Eq. 13). The dependence of protonic conductivity in poly(PFSA) membranes on water content at low water levels has been recently evaluated with molecular level modeling, based on "first principles" methods [26]. A significant result of this modeling was the elucidation that, as water molecules are added to the molecular system considered, it takes five to six water molecules per sulfonic acid group to have a clearly dissociated H_3O^+ ion around the sulfonic acid head group. At lower water contents than that, the proton is still strongly associated with the counter charge on the sulfonic acid group, as manifested by the very short calculated proton–sulfonate group distances. According to these calculations, protonic charge shielding is ineffective at less than five water molecules per sulfonic acid group, whereas around five to six waters per sulfonic acid group, water starts to provide sufficient shielding as manifested by the appearance of the well-resolved H_3O^+ group significantly further removed from the sulfonate group. This is shown in Fig. 11. One cannot fail to notice that the number of water molecules per sulfonate group required, according to these first principles calculations, for effective acid dissociation (hence significant protonic conductivity), corresponds to the transition from membrane water associated with uptake enthalpies larger than that of water condensation to membrane water of uptake energy resembling that of free water condensation [24].

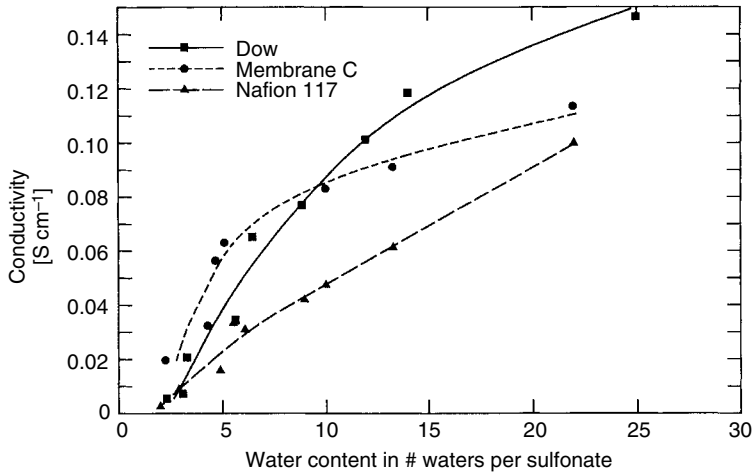


Fig. 10 Protonic conductivity versus water content for several poly(PFSA) membranes exposed to water vapor of varying activity at 30 °C [1].

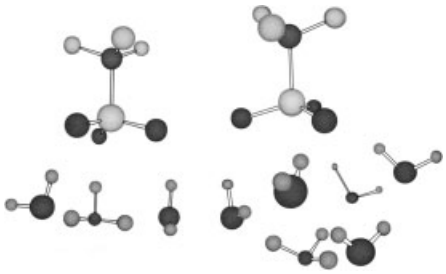


Fig. 11 First principle calculations of lowest energy structures for the triflic acid/water system. Left: $CF_3SO_3 + 3H_2O$; Right: $CF_3SO_3H + 6H_2O$. For the case of $\lambda = 3$ (left), dissociation just starts to occur, however the H_3O^+ moiety remains very close to the SO_3^- group: r(O-H) = 1.56 Angstrom. In contrast, in the case of $\lambda = 6$, separation of the hydrated proton from the conjugated anion occurs: r(O-H) jumps up from 1.74 Angstrom to 3.68 Angstrom with the sixth water molecule added [29].

Proton mobility in a poly(PFSA) membrane is measured to be very similar to water mobility in the same membranes at values of λ up to 12 [27, 28]. This strongly suggests that the mobile charged species is H_3O^+. A comparison of mobilities of protons and of water in two types of ionomers, poly(PFSA) and polyaromatic as well as in concentrated HCl solutions, is given in Figs 12 and 13. The similarity of proton mobility and water mobility can be seen to be common to different types of concentrated aqueous acid solutions, including aqueous subphases in ionomers and aqueous, concentrated HCl. This similarity is understood to result from distortions of the liquid water structure by ionic fields in such concentrated solutions, lowering the probability of proton hopping and leaving a ''vehicular mechanism'' of proton mobility, that is, proton migration as a H_3O^+ ion at an effective diffusivity similar to that of a water molecule in the same membrane. The ratio

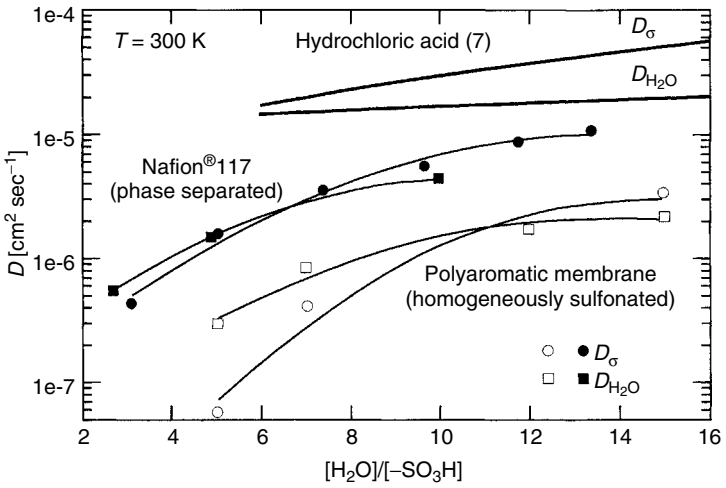


Fig. 12 Intradiffusion coefficients of water and of protons in concentrated aqueous acid solution, homogeneously sulfonated polyaromatic ionomer, and phase-separated poly(PFSA) ionomer [27].

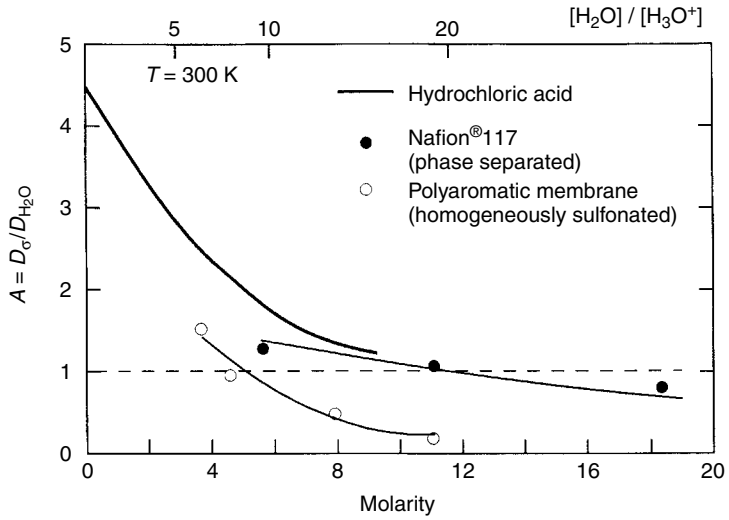


Fig. 13 The ratio of proton/water mobility in the three media described in the caption of Fig. 12.

of proton to water mobility in Nafion does increase, however, above unity, reaching 2.5 at $\lambda = 22$ [27, 28], indicating significant proton hopping in fully swollen hydrated membranes (for comparison, in bulk aqueous acid, a ratio of 4 is observed between D_{H_2O} and D_{H^+}). As can be seen from these figures, the high degree of separation (on the nanometer scale) of an aqueous subphase in poly(PFSA), is an important factor in enhancing protonic mobility in poly(PFSA) versus the case of sulfonated polyaromatic membranes of similar charge concentration.

The effect of temperature on the protonic conductivity in fully hydrated poly(PFSA) membranes can be qualitatively described as a 50% increase in the conductivity between 20 and 80 °C. However, when in contact with saturated water vapor, a rise in poly(PFSA) membrane protonic conductivity with temperature is followed by a fall above 80 °C [25], as result of the trade-off between the thermal activation of the conduction process and a decreased membrane water content with temperature when in contact with the water vapor. A continuous rise in poly(PFSA) protonic conductivity with temperature is observed for the membrane only when in contact with liquid water.

8.3.4.4 Water Currents through Ionomeric Membranes in an Operating PEFC

Sources of water and the various flux components that redistribute water in an operating PEFC are schematically shown in Fig. 14.

Water is produced at the cathode by the cell process and can enter the cell as part of (humidified) feed streams. Water is transported in the PEFC through the membrane from anode to cathode by electroosmotic drag, that is, by a water flux associated with the protonic current. The flux of water due to electroosmotic drag (moles cm^{-2} sec^{-1}) is given by:

$$N_{w,drag} = I\xi(\lambda)/F \qquad (14)$$

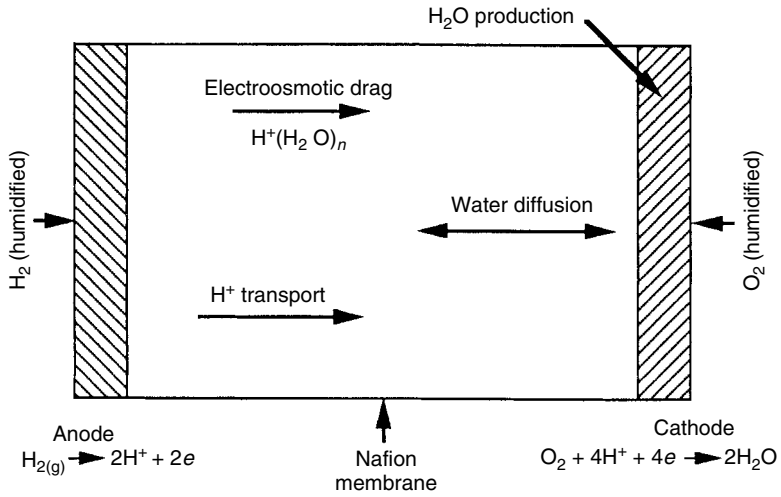


Fig. 14 Schematic representation of water sources and water fluxes that determine the water distribution profile across an operating PEFC.

where I is the current, $\xi(\lambda)$ is the electroosmotic drag coefficient at a given state of membrane hydration λ (= $(H_2O)/(-SO_3H)$ mole ratio), and F is the Faraday constant. If water supply into the anode with the fuel feed stream is insufficient, this water migratory flux can dry out the anode side of a cell under current while potentially generating an excess of water in the cathode. Buildup of water at the cathode can be relieved, in turn, by diffusion back through the membrane down the resulting water concentration gradient and by hydraulic permeation of water, for example, in differentially pressurized cells where the cathode is held at higher overall pressure than the anode. The fluxes (moles cm^{-2} sec^{-1}) brought about by the latter two mechanisms within the membrane are as follows:

$$N_{w,diff} = -D(\lambda)\Delta c/\Delta z \qquad (15)$$

$$N_{w,hyd} = k_{hyd}(\lambda)\Delta P/\Delta z \qquad (16)$$

where D is the diffusion coefficient in ionomer of water at water content λ, $\Delta c/\Delta z$ is a water concentration gradient along the thickness dimension of the membrane, k_{hyd} is the hydraulic permeability of the membrane, and $\Delta P/\Delta z$ is a pressure gradient along the membrane thickness dimension. The steady-state water profile across the ionomeric membrane for given cell current density, given anode and cathode feed humidification conditions, and given differential pressurization is a resultant of electroosmotic, diffusive, and hydraulic fluxes. In other words, a steady-state water profile in the membrane is characterized by a local steady-state water concentration gradient, $\Delta c/\Delta z$, that would generate a diffusional flux (Eq. 15) toward the anode that, together with any codirectional hydraulic flux, would buck the migratory flux of water from anode to cathode to generate a net steady state flow of water across the membrane from anode to cathode, $N_{w,net}$. At steady state, therefore, for any membrane coordinate along the thickness dimension the following holds:

$$N_{\mathrm{w,drag}} - (N_{\mathrm{w,diff}} + N_{\mathrm{w,hyd}}) = N_{\mathrm{w,net}} \quad (17)$$

In order to evaluate quantitatively the steady-state water profile in the membrane of a PEFC under given operating conditions, the membrane transport properties required can be seen to include the following: (1) water uptake by the membrane as function of water activity in contact with the major membrane surfaces, $\lambda(a_{\mathrm{W}})$, (2) the diffusion coefficient of water in the membrane as a function of membrane water content, $D(\lambda)$, (3) the electroosmotic drag coefficient as a function of membrane water content, $\xi(\lambda)$, and (4) the membrane hydraulic permeability, $k_{\mathrm{hyd}}(\lambda)$, the latter being relevant when hydraulic pressure difference exists across the membrane.

8.3.4.4.1 **Diffusion of Water in poly[PFSA] Membranes** To describe diffusion of water through the membrane in the presence of a water activity gradient, an appropriate interdiffusion coefficient must be determined. Experimental methods used to study diffusion of water in these polymers, such as radiotracer and pulsed gradient spin-echo NMR techniques, probe intradiffusion coefficients, often referred to as tracer or self-diffusion coefficients, determined in the absence of a chemical potential gradient. Intra- and interdiffusion coefficients are related for the case of diffusion of a small molecule in a polymeric matrix as follows [28]:

$$D_{\mathrm{inter}} = D_{\mathrm{intra}}\{\mathrm{d}(\log a)/\mathrm{d}(\log C)\} \quad (18)$$

Reference 28 reports pulsed field gradient NMR studies of water motion in Nafion 1100, yielding intradiffusion coefficients close to 1×10^{-5} cm^2 sec^{-1} for samples in contact with saturated vapor. For liquid-immersed Nafion membranes, other NMR diffusion imaging studies [29] also showed diffusion coefficients similar to those measured in free, liquid water -2.2×10^{-5} cm^2 sec^{-1} in liquid water versus 1.7×10^{-5} cm^2 sec^{-1} in the fully hydrated membrane. To convert the intra-(self)diffusion coefficients (D_{self}) to inter-(Fickian) diffusion coefficients (D_{chem}) [28], measured self-diffusion coefficients have to be corrected for the variation of the water activity coefficient with water content (as well as for effects of swelling of the polymer). The resulting D_{chem} for water in the Nafion membrane varies between 2×10^{-6}–6×10^{-6} cm^2 sec^{-1} at 30 °C for water contents between $\lambda = 2$ and 14. The same authors subsequently carried out measurements on various membranes at 80 °C. These results are summarized in Fig. 15. The approach of calculating the Fickian diffusion coefficient from self-diffusion coefficients measured, for example, with NMR and the measured dependence of water content on water activity to derive $d(\log a)/d(\log C)$ (Eq. 14) has been further validated in later work by Motupaly et al. [30].

Water transport in the ionomeric membrane driven by a water population gradient, described first in terms of water diffusion (e.g., Ref. 28), was subsequently modeled using liquid flow in a porous medium as an alternative description [31]. The latter description projected higher water flow rates under a given water population gradient as compared with [28] (and, consequently, more flat water profiles in the membrane were projected). To a large degree, the apparent disagreement resulted from the ''diffusion model'' described in Ref. 28, having covered the case of a membrane equilibrated with water vapor, whereas the ''liquid flow in a porous medium'' model [31] covered the case of a membrane containing liquid water, that is, equilibrated, at least on the

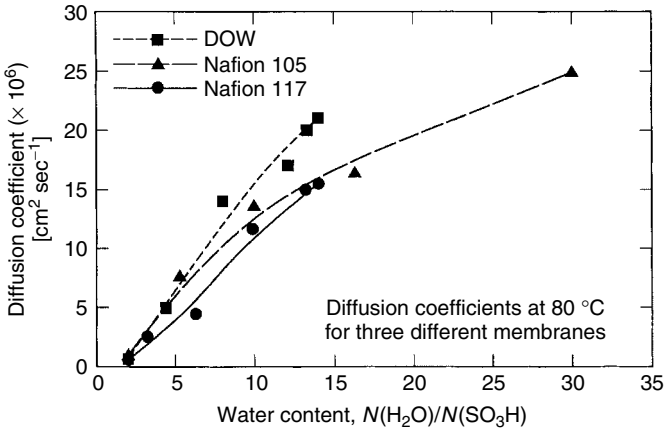


Fig. 15 ^{1}H diffusion coefficients for poly(PFSA) membranes as function of water content at 80 °C [1].

cathode side, with liquid water. The significant difference in water content and in water transport properties between a membrane in contact with water vapor and a membrane in contact with liquid water is discussed in Sect. 8.3.4.2.1.

Reference 32 reports an activation energy for water transport in Nafion on the order of 0.17 eV (3.9 kcal mol^{-1}) for water contents ranging from $\lambda = 8 - 14$. At lower water content examined, $\lambda = 3.5$, an activation energy of 0.22 eV (5.1 kcal mol^{-1}) is reported over the temperature range 20–90 °C. The increase in activation energy for water motion with decreasing water content [32] accords well with the notion that water mobility is increasingly hindered at low water contents by shrunken pores and stronger electrostatic interactions. On the basis of these numbers for the activation energy, an increase by about factor 2 between 20 and 80 °C is expected in the water diffusion coefficient in poly(PFSA) membranes of higher water contents, and by about factor 4 at the lowest water content.

8.3.4.4.2 **Electroosmosis** The number of water molecules, $\xi(\lambda)$, carried per proton moving across an ionomeric membrane of water content λ uniform is the so-called electroosmotic drag coefficient. When $\xi = 2$, for example, two water molecules are carried across the membrane per single protonic charge moving from anode to cathode under conditions ensuring that the protonic current is the only driver of water flux. The electroosmotic drag coefficient is a key parameter in the determination of the water profiles across the membrane in an operating PEFC. This proton-coupled drag of water is a major reason for the need of effective water management in PEFCs, as well as in DMFCs using liquid, water/methanol feed streams. Note the important difference between the *electroosmotic drag coefficient,* $\xi(\lambda)$, a characteristic of an ionomeric membrane of some fixed water content, measured under conditions of flat membrane water profile, and the *net water flux per proton through an operating fuel cell*. The latter is the resultant of all cell water transport components (see Fig. 14 and Eq. 17)

and of the boundary conditions defined by the water content of the feed streams.

A recent comprehensive measurement of drag coefficients in ionomeric membranes fully equilibrated with liquid water has been based on monitoring the water flux through the cathode exhaust in a DMFC with aqueous, dilute methanol feed [33]. Above some cell current level, this specific system allows to establish a flat water profile in the membrane and, consequently, the measured water flowrates across the membrane become associated exclusively with the protonic current. Results of this study are shown in Fig. 16. The study reconfirmed the water drag coefficient measured earlier at ambient temperature for the protonic form of poly(PFSA) membranes, that is, $2.5H_2O/SO_3H$ and showed further that, as long as the membrane is fully equilibrated with liquid water, this coefficient would increase to $5H_2O/SO_3H$ at a temperature of 80 °C.

An elegant solution to the problem of obtaining electroosmotic drag coefficients in ionomeric membranes under conditions of *vapor phase* equilibration is described in Ref. 34. The EMF method described involves measurements of the potential across a membrane sample exposed on each end to different water activities. The potential difference $\Delta\Phi$ is determined by $\xi(\lambda)$ according to [34]:

$$F\Delta\Phi = \xi(\lambda)\{RT\log(a_{\text{water, right}}/a_{\text{water, left}})\} \tag{19}$$

Reference 34 reported a drag coefficient of protons in hydrated Nafion 117 membranes, which is essentially constant at $1.4H_2O/H^+$, in the range $\lambda = 5-14$, and gradually drops to zero between $\lambda = 5$ and 0. Others [29] have subsequently adopted the Fuller–Newman EMF technique and concluded from their data a constant water drag of roughly $1.0H_2O/H^+$ over a large range of water contents in vapor

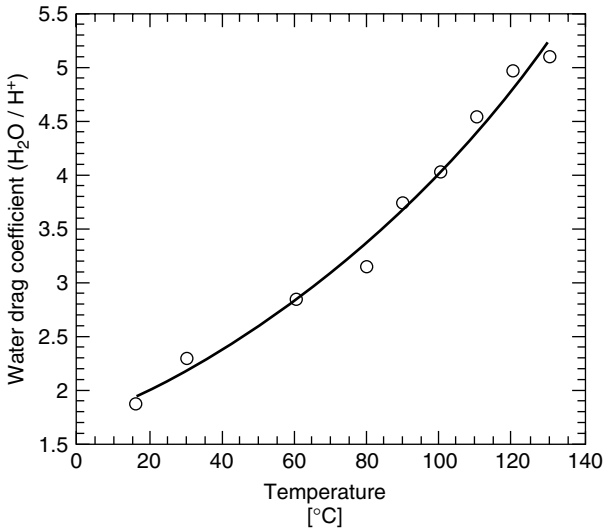


Fig. 16 Water drag coefficient as function of temperature in a direct methanol fuel cell with aqueous, liquid anode feed [33].

equilibrated membranes, from $\lambda = 14$ down to $\lambda = 1.4$. Further recent measurements [35] show that several poly(PFSA) membranes exhibit similar electroosmotic behavior, that is, a drag coefficient of close to $1.0 H_2O/H^+$ over a wide range of water contents for a membrane equilibrated with vapor phase water. The lack of dependence of the drag coefficient in such membranes on details of membrane chemistry, or nanostructure, suggests that the drag coefficient is determined by the nanometer scale of the membrane porosity and the significant deviation from bulk water properties in the surface-charged nanopores of the membrane. These features prevent much larger water fluxes per proton, as measured in electrokinetic (EK) pumping through micro-(rather than nano) porous membranes that contain a high fraction of free, bulk like water.

8.3.4.4.3 **Modeling of the Water Profile and Net Water Flow Rate across the Ionomeric Membrane under Current** The steady-state water profile in a PEFC under load corresponds to a constant rate of water flow through the thickness dimension of the cell, set, in turn, by the cell current, the water transport properties of the membrane and other cell components, and by boundary conditions defined by the degree of humidification of the fuel and air feed streams. The steady-state water profile across the thickness dimension of the membrane has a direct impact on cell performance. Maintenance of high water content across the membrane under the highest cell current densities is required to ensure high membrane protonic conductivity through the cell dynamic range targeted. A mathematical description of water transport in the membrane of a PEFC under load requires as input the dependencies on membrane water content of the electroosmotic drag coefficient, $\xi(\lambda)$, and the water diffusion coefficient, $D(\lambda)$. Experimental evaluation of these parameters as a function of λ was described in sections 8.3.4.4.1 and 8.3.4.4.2. In addition, boundary conditions at the anode and the cathode side of the membrane have to be defined for some given cell operation conditions, specifying water levels in the fuel and air feed streams. Experimentally derived diffusion and electroosmotic drag coefficients of water in Nafion as function of membrane water content were first used in Ref. 36 to model the steady-state water profile and the resulting protonic conductivity in the membrane of an operating PEFC. Water flux through the electrodes was assumed in this model to take place in the gas phase (water vapor) only.

The experimentally determined isotherm for water sorption by the membrane (see Fig. 9) was used to convert from water vapor activity in the gas phase at the interface to water content in the membrane at the surface and considering generation of water at the cathode, the water level in the membrane surface adjacent the cathode was assumed accordingly to be $\lambda = 14$ (or somewhat above, see Ref. 36). The components of water flow considered within the membrane included electroosmotic drag and diffusion, whereas hydraulic flux was considered insignificant (i.e., very small hydraulic pressure drop across the membrane) [36]. Consequently, the basic equation describing the flux within the membrane becomes

$$N_{w,net} = \alpha(I) = (I)\xi_{\lambda=22}(\lambda/22) - (\rho_{dry}\tilde{E}\, W\, D_\lambda(d\lambda/dz) \quad (20)$$

where $N_{w,net}$ is the net steady-state flux of water through the membrane from anode to cathode when the current density is I,

$\xi(\lambda)$ is the number of water molecules carried per proton migrating through the membrane at a local water level λ, and D_λ is the diffusion coefficient at the same local water level. The factor $(\rho_{\text{dry}} \tilde{E} \dot{W})$ converts gradients in water content, λ, in the membrane to gradients in water concentration in moles per cubic centimeter: ρ_{dry} is the weight density of the dry membrane and EW is its EW. The water flux component $\alpha(I)$ in the r.h.s. of Eq. (20) describes an anode boundary condition where a flux of water proportional to the cell current (hence to the humidified anode feed rate) is assumed to enter the membrane from the anode. As can be seen from Eq. (21), the detailed dependence of $\dot{\xi}$ on $\dot{\lambda}$, as described in the preceding section, is approximated here by a linear relationship between $\dot{\xi}$ and $\dot{\lambda}$, with ξ varying between its measured value, $\xi_{\lambda=22}$ at full membrane hydration ($\lambda = 22$) and zero at $\lambda = 0$.

From the last equation, the local gradient in λ in the membrane could be expressed as follows:

$$d\lambda/dz = [2\xi_{\lambda=22}(\lambda/22) + \alpha] \times [I(EW\rho_{\text{dry}} D_\lambda(\lambda)] \qquad (21)$$

and Eq. (21) can be integrated numerically so as to satisfy the boundary conditions at the anode and cathode/membrane interfaces (the latter being $\lambda = 14-22$, depending on whether water vapor only or vapor + liquid are present at the cathode/membrane interface).

Solved profiles of the water level λ across the membrane in an operating PEFC are shown in Figs 17 and 18. Figure 17 shows the steady-state profile of water in a Nafion 117 membrane (175 µm thick), as calculated with this model at four different current densities. The calculated profile of water is seen to become steeper as the PEFC current density increases, resulting in a significant depletion of water in the membrane near the anode (in spite of the water vapor–saturated hydrogen feed stream assumed). When the water level through part of the membrane thickness drops as low as Fig. 17 projects for the anode side of the membrane under higher current, the membrane resistance rises substantially (Sect. 8.3.4.3). A first-order verification of the water profile calculated can be done by comparing the calculated resistance across the membrane corresponding to the calculated profile to the measured high-frequency resistance of the fuel cell. As Fig. 18 shows, calculations of the resistance across the membrane from the water profiles derived predict a $\times 5$ rise in membrane resistance for a 175-µm thick Nafion membrane (Nafion 117) when the current through the membrane reaches 1 A cm^{-2}. At the same time, as Fig. 18 shows, Nafion membranes of the same EW(1100), which are 100 µm thick, or thinner, are projected to see a much smaller problem of water depletion near the anode at high cell currents. The fundamental reason for the ability of such thinner membranes to maintain a relatively flat water profile up to high current densities is that the rate of diffusional backflow form cathode to anode per given gradient in water population (set by the electroosmotic drag) is inversely proportional to membrane thickness. Figure 18 projects that it should be sufficient to lower the Nafion membrane thickness from 175 to 50 µm to completely eliminate any increase in membrane resistance up to a cell current density of 2 A cm^{-2}. Results of measurements of high-frequency resistance in PEFCs with membranes of different thicknesses as function of current density fully confirm the predictions of these model calculations (see, e.g., Fig. 8).

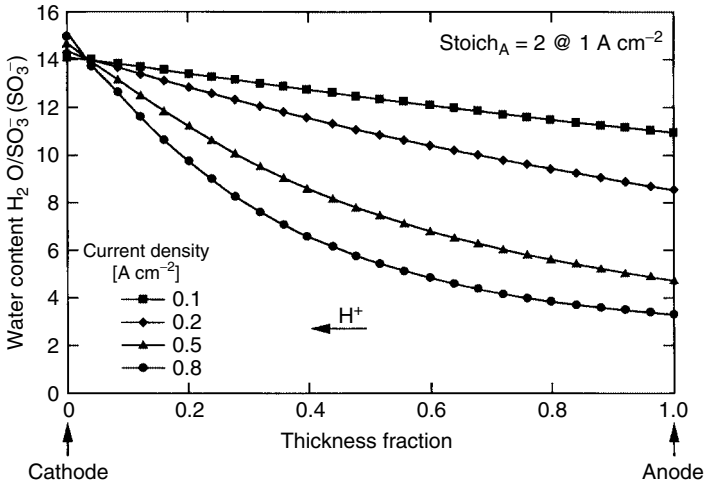


Fig. 17 Calculated water profiles in a Nafion™ membrane in an operating PEFC as functions of current density. Cell and gas humidifiers are assumed to be at 80 °C [36].

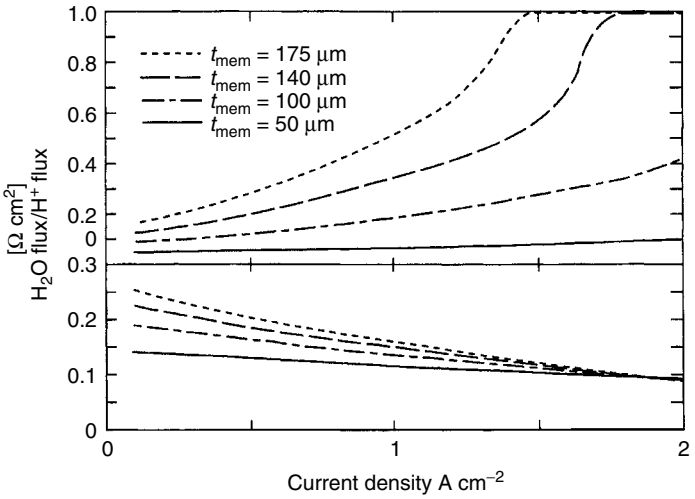


Fig. 18 Computed R_{cell} (upper panel) and the ratio water flux/proton flux through the cell (lower panel) under the conditions described for Fig. 17 [36].

Correspondingly, in state-of-the-art PEFC stacks, the poly(PFSA) membranes employed are typically 20–50 µm thick and, while requiring more care in handling and in stack engineering, such very thin membranes are very effective in facilitating full membrane hydration up to very high cell currents. In addition to the calculated high-frequency resistance of the cell (top part), Fig. 18 also shows (bottom

part) calculated net fluxes of water, $N_{w,net}$, through the membrane for membranes with identical transport properties differing only in thickness.

The model predictions described in the preceding text for water distribution across the thickness dimension of the PEFC membrane (Figs 17 and 18) have recently been directly confirmed in an elegant experiment providing MRI images of water distribution across the PEFC membrane thickness dimension in cells under current [37]. Images of water in a relatively thick (340 μm) PEFC membrane, recorded at zero cell current, at low cell current density and a higher cell current density, are shown in Fig. 18, under supply to the PEFC of dry oxygen and dry hydrogen gases. The highest local water level, $\lambda = 6.5$, seen here in the membrane imaged, is relatively low because of the dry gases supplied. The image clearly shows, however, the steady-state gradient between higher water content at the cathode and a lower water content near the anode, becoming stronger with increase in cell current as predicted by the membrane water transport model for such thick ionomeric membranes (Fig. 17).

Further modeling work done more recently reveals expected variations of both temperature and water content in a PEFC, together with the associated variations in protonic conductivity along the gas flow channel, that is, along the active surface area of the MEA. Such lateral distributions are evaluated and discussed in Ref. 38.

Measurements of the net water flux across the PEFC (e.g., Ref. 36) typically reveal small net water fluxes across the cell per unit charge passed, similar to those calculated by the above model – see bottom part of Fig. 18. Clearly, these net water fluxes across the cell are significantly smaller than the electroosmotic driven water flux in the membrane, as expected from the opposing water back-diffusion under steady-state conditions (see Eq. 20). This result is typical, however, to an anode feed stream with water in vapor form only and absence of hydraulic pressure drop across the membrane. Both conditions were assumed in deriving the profiles in Fig. 16. Possible flow of liquid water through the cell under hydraulic pressure drop was treated elsewhere as part of a detailed mathematical model for the PEFC [39] and results based on measured hydraulic permeabilities for both the membrane and the electrodes were described. Water transport in liquid form through the electrode (basically through the electrode backing layer) was considered again in a recent modeling effort [40] targeting quantitative evaluation of the net, through-the-cell water flux, measured as a function of the cell operation conditions. Measured water permeability of cell electrodes helped explain observed net water transport rates through the cell with saturated water vapor in the anode, such rates ranging between +0.21 water molecules per proton (for cell with Nafion 105 membrane) and −0.09 water molecules per proton (for cell with Nafion 112 membrane). Positive flux corresponds here to net water flow from anode to cathode and negative flux to net water flow from cathode to anode. In all cases of cells tested on the basis of the Nafion 105 membrane, the measured net flux of water through the cell per proton could be shown to closely correspond to the rate of water supply to the anode side of the membrane, as determined by the flux of vapor through the anode under vapor-saturated gas conditions [40]. In other words, the rate of electroosmotic drag of water from anode to cathode

adjusts according to anode humidification conditions, such that it exceeds the rate of water back-diffusion by the rate of water supply to the membrane through the anode (see Eq. 17 with $N_{w,net} = N_{w,across\ anode}$).

8.3.4.4.4 **The PEFC Ionomeric Membrane – State of the Art** Poly(PFSA) membranes have been to date the membrane of choice for PEFCs as well as for DMFCs. Poly(PFSA) membranes exhibit not only high protonic conductivity, typically, 0.2 S cm^{-1} in the fully hydrated membrane at typical cell operation temperature, but also high water transport rates (diffusion coefficients of 2×10^{-5} cm^2 sec^{-1} for the fully hydrated membrane), as required for minimized ohmic losses in the membrane. A 50-μm-thick poly(PFSA) membrane should contribute only 0.025 Ω cm^2 to cell resistance (voltage loss of only 25 mV at 1 A cm^{-2}) at 80 °C, as long as the water level is sustained at full hydration through the thickness dimension of the membrane in the operating cell. The latter requirement is facilitated by the high diffusivity of water in these membranes and, indeed, the overall series resistance measured in PEFCs with such poly(PFSA) membranes is typically as low as 0.05 Ω cm^2, the excess over the 0.025 Ω cm^2 contributed by various contact resistances in the cell. Watanabe et al. [113] described enhanced humidification of PEFC membranes, achieved by seeding the membrane with Pt nuclei to encourage water generation within the membrane by oxygen/hydrogen recombination. Subsequent similar efforts have been described, where micro or nanoparticles of ceramic materials were used for a similar purpose. Such efforts are probably mostly relevant in the pursuit of the coveted ''high-temperature polymer membrane electrolyte'' because effective hydration of thin poly(PFSA) membranes in the operating cell is readily achieved with present membrane technology under 100 °C. The challenge of a ''water-balanced PEFC'' can be addressed with the so-called enthalpy exchanger, practiced quite extensively in PEFC systems [6]. In such exchanger, the inlet stream of ambient air passes in counterflow to the highly humidified and warm air exhaust stream, with Nafion, or similar membrane separating between the two streams. In this way, the air inlet feed stream is both warmed up and humidified very effectively prior to entering the cell cathode, the cell product water using the stack generated heat and. Such solutions, provided at the membrane thickness composition and at the system level, respectively, have paved the way for PEFC operation at or below 80 °C with no humidification by water from external sources required.

The state of the art of membrane technology and related water-management solutions provided to-date seem to have established a good basis for high-performance PEFC stacks. However, remaining unanswered requirements from the poly(PFSA) membrane seem still significant. They include primarily sufficient long-term stability in hydrogen/air cells, particularly in light of vulnerability to attack by $OH^{\bullet}$ radicals and hydrogen peroxide. For the latter reason, for example, crossover rates of hydrogen and oxygen gas through such fuel-cell membranes, which were considered sufficiently low from just the fuel loss perspective (the cross-membrane gas fluxes are limited to just several milliamperes per square centimeter), are now understood to be the route of OH radical formation by hydrogen/oxygen recombination at one of the catalyzed electrodes. Another weakness is structural volunarability

under dry/wet cycles which bring about large dimensional changes in poly[PFSA] membranes. And finally, the sensitivity of poly(PFSA) conductivity to membrane water content remains a fundamental drawback in limiting the operation of such polymer electrolyte PEFCs to temperatures below 100 °C. Alternative membranes for operation above 100 °C have been prepared and tested; however, the result to date has always been cell performance significantly lower than that of a PEFC employing a poly(PFSA) membrane because the potential for higher activity rendered by higher cell temperatures is upset by the low activity of PEFC electrode processes at low interfacial water content (see section 8.3.7.2.2).

8.3.5 From Ionomeric Membrane to the Membrane-electrode Assembly

Given a high-quality poly(PFSA) membrane of properties described in the last section, assembling an electrochemical cell based on such membrane requires to define the nature of the catalytic electrodes and the method(s) of coupling them effectively to the membrane on one side and the reactant distributor/current collector on the other side. A significant challenge is presented in PEFC and DMFC technology by the demand of high catalytic activity that needs to be achieved at relatively low cell temperatures. An additional important challenge in polymer electrolyte systems has to do with the ''pseudo–solid state'' of the polymeric membrane, in other words, the absence of any liquid electrolyte. The latter presents a special need for the effective intermixing of the ion-conducting, ionomeric component with a carbon-supported metal alloy catalyst, to form a catalyst layer of mixed conductivity and of uniform catalyst distribution, and have such layer very well bonded to the ionomer membrane surface. Three types of tools are available, in principle, to answer this overall challenge: optimized metal alloy catalyst surface properties, maximized catalyst dispersion and effective catalyst layer composition and structure that enables high catalyst surface utilization.

Optimization of catalyst composition and preparation/fabrication for the PEFC or DMFC cathode has been pursued during the last decade with one significant success being the replacement of Pt by some Pt alloy catalysts, including PtNi and PtCr [41]. By the mid-2000s, PtCo alloys eventually emerged as the cathode catalysts of choice for PEFCs. The latter binary cathode catalysts have been reported to provide intrinsic catalytic activity in oxygen reduction in a PEFC cathode that is higher by a factor of 3–5 versus that of a non-alloyed Pt catalyst of similar Pt mass loading. This is an advantage of high value, particularly in PEFC applications of highest sensitivity to the price of platinum. Catalyst composition optimization at the anode side of the PEFC and DMFC has been primarily centered on the optimization of the composition and the refined preparation of PtRu alloy catalysts. The latter exhibit significant tolerance to 10–100 ppm CO in hydrogen-fueled PEFCs and high activity in electrooxidation of methanol in the DMFC anode. These electrocatalysis R&D issues are discussed in further detail in Sects 8.3.7.1 and 8.3.8.1, respectively.

Perhaps the most important advancements in electrocatalysis that contributed mostly to the development of low-temperature fuel-cell technology belong to the domain of ''practical'' electrocatalysis and include maximized catalyst dispersion and effective ''packaging'' of the required

surface area into thin catalyst layers of optimized composition structure, and of good bonding to the substrate surface. Highly dispersed Pt or Pt alloy catalysts can be prepared using a variety of techniques. The most common is the formation the Pt colloidal particles from solution onto partly graphitized carbon powder (e.g., Vulcan XC-72), with reduction by various possible reducing agents taking place during and following the formation of such colloidal particles. The final product is typically metal particles of the order of 2 nm in diameter, supported onto carbon power of particle size around 10 nm. For catalysts consisting of spherical particles of diameter $2r$, application of catalyst mass m per square centimeter of membrane area results in overall catalyst surface area, a_{cat}, per square centimeter of membrane area of

$$a_{cat} = 3m/r\rho \qquad (22)$$

where ρ is the mass density of the catalyst. For example, at $m = 0.4$ mg Pt/cm^2 of membrane area, $r = 1$ nm, and ρ (for Pt) of 20 g cm^{-3}, a_{cat} is calculated as 600 cm^2 Pt/cm^2 of membrane area. Such highly dispersed catalysts thus allow, in principle, to reach a factor 600 ''amplification'' of the geometric surface area of the electrode with mass loading of only 0.4 mg Pt/cm^2 of electrode area. Such a high level of catalyst dispersion has been one of the most important keys to the high power output and conversion efficiencies achieved in PEFCs.

Equation (22) shows that the total surface area of the catalyst at given level of dispersion (given r) increases linearly with overall catalyst mass loading, m. The same is not true, however, for the *effective* or utilizable catalyst surface area. It has been the experience of developers of PEFC and DMFC MEAs that, catalyst utilization, as a rule, drops with increase in catalyst layer thickness. Typically, catalyst activity increases linearly with loading (and, therefore, with catalyst layer thickness), up to an overall catalyst layer thickness of 4–5 μm, but, for thicker layers of higher catalyst loadings, the incremental activity increase with incremental increase in loading starts to drop. An activity plateau is typically hit at catalyst layer thicknesses somewhat above 10 μm, with further increase in loading (i.e., in catalyst layer thickness) resulting eventually in a drop of electrode activity. The reasons for such a drop in catalyst utilization at higher catalyst layer thicknesses have been clarified by detailed modeling of the critical coupling between the interfacial process at catalyst sites distributed evenly in the volume of the catalyst layer and the transport processes along the catalyst layer thickness dimension of reactant, product, and protonic and electronic charges. Sufficient rates of all the above are simultaneously required to ensure high catalyst utilization [13], limiting the activity of sites located too far from the ionomeric membrane surface because of proton transport limitations and of sites placed too far from the reactant flow channel because of gas transport limitations within the catalyst layer [13]. Figure 19 provides a quantitative elucidation of the effects of limited transport rates on catalyst utilization in a thin catalyst layer with uniform volume distribution of catalyst particles. The figure depicts results of modeling of distributions along the normalized thickness dimension of a 5-μm-thick PEFC anode catalyst layer, of anode potential (top), hydrogen reactant concentration (middle), and current generation (bottom). Two cases have been considered for comparison: a 100% hydrogen anode feed and a 10% hydrogen feed (both dry gas composition) [42]. The figure

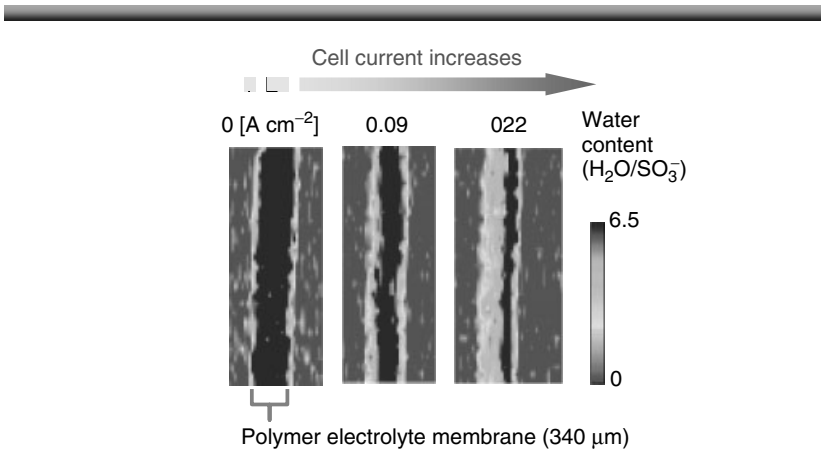


Fig. 19 MRI of water distribution in a PEFC at zero current and under two current densities. The cathode is on the right side of the MEA imaged [37].

shows that, as result of limited rate of reactant diffusion into the catalyst layer in the case of the 10% hydrogen feed, the concentration of the reactant and, consequently, the current production, are both calculated to be very low at fractional widths of the catalyst layer larger than 0.5, away from the catalyst layer/gas interface. On the other hand, if the effective protonic conductivity of the catalyst layer is an order of magnitude lower than that of the neat ionomer, a significant drop in overpotential across the catalyst layer could make the half of the catalyst layer away from the membrane significantly less effective in current production. This is seen to be the case for the combination of low protonic conductivity and high reactant concentration. It can be readily understood from Fig. 20 why, at given transport characteristics typical for composite, Pt/C//ionomer catalyst layers, utilization of the catalyst could strongly drop when the thickness of the layer exceeds 5–10 μm, particularly in the case of low reactant concentration. In contrast, thicker layers could still enhance electrode performance in the case of higher reactant concentrations, provided the protonic (and electronic) conductivity can be maintained at a sufficiently high level in the composite structure while maintaining, at the same time, effective reactant access through the layer's void volume. One unfortunate consequence of such transport limitations within the catalyst layer is that catalysts of much lower cost but intrinsically low activity (significantly inferior to that of Pt) cannot be effectively utilized by applying higher mass loadings, which could be perfectly affordable in such cases. Transport limitations of the type shown in Fig. 20 would not allow effective catalyst utilization through the much thicker catalyst layers needed to contain the much higher total surface area of a catalyst of low specific activity, that would be required to generate a certain areal current density demand.

Figure 21 shows a TEM micrograph of a carbon/supported Pt catalyst with Pt catalyst particle size centered around 2 nm. The micrograph depicts the nanogeometry of such supported Pt catalysts and

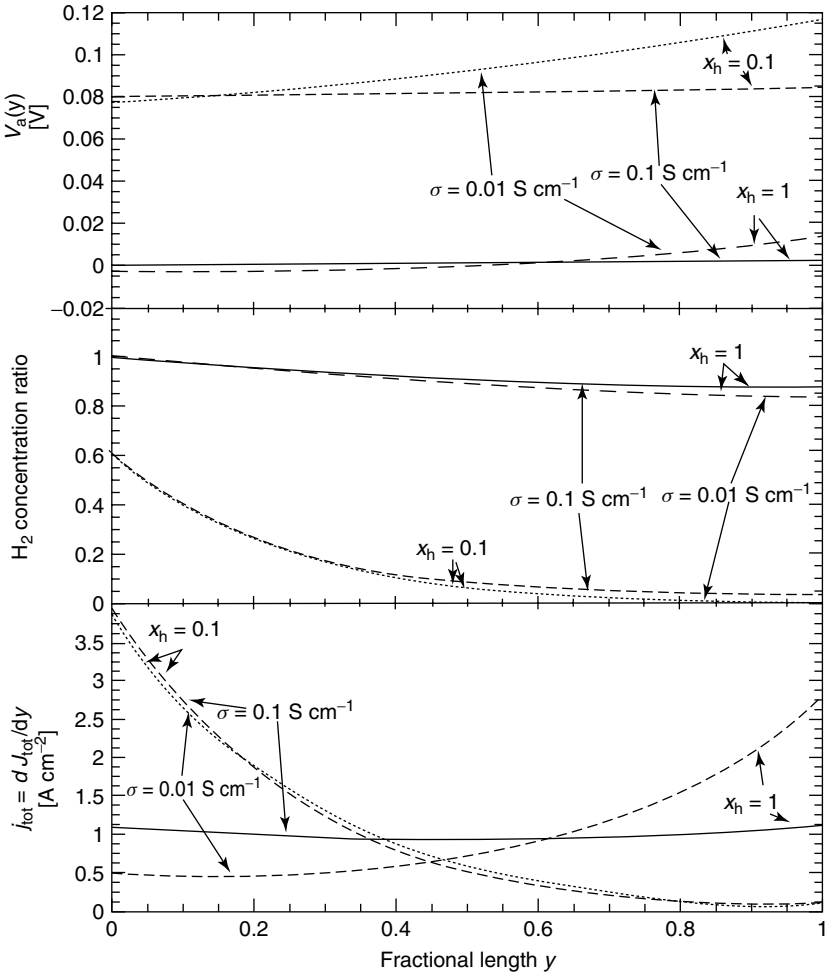


Fig. 20 Distributions at current density of 1 A cm^{-2}, of electrode potential (top), reactant concentration (middle), and current generation (bottom) in a PEFC anode catalyst layer 5 μm thick, as result of limited transport rate of the hydrogen gas reactant and/or the limited transport rate of protons. Two cases of reactant concentration, 100% hydrogen and 10% hydrogen in the dry gas and two cases of effective protonic conductivity in the catalyst layer, 0.1 and 0.01 S cm^{-1}, are considered in these calculations. A value of 2×10^{-4} cm^2 sec^{-1} was used as estimate for effective D_{H_2} in the catalyst layer.

clarifies the challenge of utilizing a high fraction of the total Pt surface given in this form. High catalyst utilization requires to secure maximum transport rates of reactant, protons and electrons in composite Pt/C//recast ionomer layers (of significant void volume) containing these dispersed catalysts. Consequently, optimized preparation of the Pt catalyst/recast ionomer composite and its proper application to

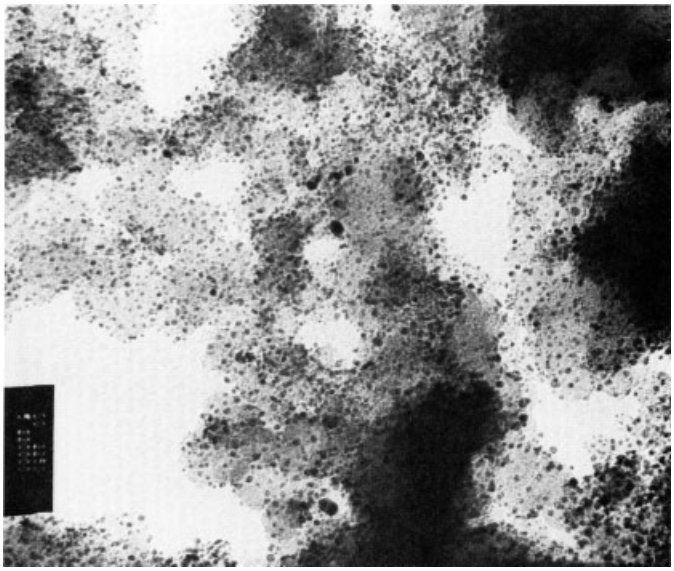


Fig. 21 Electron microscopy image of Pt/C catalyst with average Pt particle diameter of 2 nm.

the membrane are key to catalytic activity high in the PEFC and DMFC. A polymeric membrane with catalyst layers applied directly to its two major surfaces has been referred to as the catalyst-coated membrane (CCH) or the membrane/electrode assembly (HEA). In other modes of preparation of the catalyst layer in PEFCs and DMFCs, the catalyst layer is applied onto the carbon cloth or paper backing, and the catalyzed carbon paper is hot pressed, in turn, to the polymeric membrane. The specific mode of preparation of the catalyst layer could further vary within these two general application approaches, as summarized in Table 3.

As of year 2005, methods (A3) and (B1) are mostly practiced, with (A4) emerging as an advanced concept targeting CCMs with ultralow catalyst loadings.

A variety of techniques for the preparation of catalyst layers in PEFCs is given next in their chronological order of development. The earliest method of preparation of MEAs [43] described catalytically active layers prepared from finely divided metal powders (blacks) mixed with PTFE as binder. The drawback of this mode of preparation of catalyst layers is the low catalyst utilization caused by lack of efficient access of protons to a large fraction of catalyst particles located away from the membrane electrolyte and lack of efficient gas access to another fraction of catalyst sites deeply embedded in the polymer. The result of limited dispersion and low utilization in such Pt black/PTFE mixtures was the requirement of very high Pt loading, typically 4 mg Pt/cm^2 (geo.) to obtain satisfactory cathode performance, particularly in operation on air [44]. Another approach to the direct deposition of Pt onto the membrane surface was adopted by several workers in exploratory studies of electrochemical processes at the

Tab. 3 Modes of catalyst layer preparation and application

(A) Bonding to the membrane	*(B) Bonding to carbon cloth/paper substrate*
(A1) Applied and hot-pressed Pt /PTFE	(B1) Pt/C//PTFE applied to substrate and impregnated with recast ionomer
(A2) Electroless deposition of Pt on membrane	(B2): (B1) + sputtered Pt layer
(A3) Applied and hot-pressed Pt/C(or Pt)//ionomer within ionomer-impregnated carbon paper	(B3) Pt catalyst electrodeposited
(A4) Alloy catalyst sputtered onto array of nanowhiskers and the catalyzed array embedded into the membrane surface	

Pt/ionomer membrane interface [45–47]. The latter approach was based on various electroless deposition techniques for the application of a film of Pt to the surface of an ionomeric membrane. The original method suggested [45] was based on exposure of one side of the membrane to an anionic salt of the metal ($PtCl_6^{-2}$) and the other side to a reducing agent (e.g., N_2H_4). The latter diffuses through the membrane to react at the surface adjacent to the metal salt solution and form a Pt film. While simple and effective in exploratory studies, this approach yields relatively large Pt particle sizes and cannot allow good simultaneous access of gas, electrons, and protons to most of the surface of the metal particles deposited.

The first move toward substantial lowering of Pt loading employed in PEFCs, while maintaining comparable cell performance, was described by Raistrick [7]. Raistrick experimented with gas-diffusion electrodes developed for PAFCs. In such electrodes, the catalyst layer is a mixture of carbon-supported platinum (Pt/C) and PTFE, deposited onto a carbon cloth or paper. This is done usually by filling the macropores first with a mixture of (uncatalyzed) carbon powder and PTFE, with the catalyst layer subsequently being deposited onto one surface of the pretreated carbon cloth/paper. Compared to catalysts based on Pt black, an important advantage of the carbon-supported Pt catalyst was the higher degree of Pt dispersion. Pt particle sizes of 2 nm in diameter are easily obtained in the carbon-supported form by various protocols described in the patent literature. This is compared with a typical particle size of about 10 nm obtainable in Pt blacks. TEM images of carbon (Vulcan X-72)-supported Pt catalyst of 20% Pt by weight are given in Fig. 21. Trying to take advantage of the higher dispersion of the Pt/C catalyst, Raistrick started his investigation by hot-pressing gas-diffusion electrodes of this type directly to an ionomeric membrane and realized that the poor cell performance obtained consistently was caused by lack of protonic access to the majority of catalyst sites not in intimate contact with the membrane. Effective protonic access could not be met within the 50-μm-thick Pt/C//PTFE catalyst layer because Pt sites distributed in a catalyst layer bonded by PTFE are ''protonically disconnected'' from the ionomeric membrane. To overcome this critical limitation, Raistrick impregnated the Pt/C//PTFE catalyst layer

on the carbon cloth with a solution of Nafion®, thereby adding to the catalyst an ion-conducting component prior to hot-pressing the impregnated electrode onto the membrane [7]. This step of catalyst layer impregnation with recast ionomer resulted in a very significant increase in performance and demonstrated, for the first time, significant PEFC performance with Pt loading as low as 0.4 mg cm^{-2}, that is, an order of magnitude lower than the loadings employed in the GE/Hamilton Standard space PEFC technology. This first result of Raistrick with impregnated gas-diffusion electrodes [7] was a breakthrough in PEFC technology in two ways. Firstly, it demonstrated that PEFCs could be effectively operated with Pt/C catalysts of much lower Pt loadings, thus bringing the loading and cost of precious metal catalyst per unit power output down by an order of magnitude. In addition, and perhaps even more important for further developments to follow, this experiment revealed the criticality of ''triple access'' (gas, electrons, protons) to each active catalyst site and with that opened the way to the further optimization of PEFC catalyst layers of the desired combination of low loadings and high performance. Further significant improvements of the initial idea described by Raistrick were achieved by Srinivasan and coworkers. They described successful attempts to increase the performance of impregnated Pt/C layers on carbon support hot pressed to the membrane by increasing the temperature and gas pressure [48], optimizing the percentage of Nafion impregnant [8], replacing the 10% Pt/C 100-μm layer used in the PA electrode technology with a 20% Pt/C 50-μm layer [49], and adding a thin film of sputtered Pt on the front surface of the carbon electrode prior to impregnation [50]. The last two modifications shorten the average distance traveled by protons to access a Pt catalyst site and generate, particularly in the case of the sputtered film, a significant number of active Pt sites very close to the membrane.

An important factor that can limit performance in PEFCs prepared with such catalyzed gas-diffusion electrodes is the nature of the bonding established between the catalyzed carbon electrode and the membrane surface. Conditions of hot-pressing applied to achieve such a bonding, in terms of temperature, pressure, and hot-press duration, have to be optimized for each ionomeric membrane. During hot-press, the glass-transition temperature of the ionomer (typically, 130–160 °C) is exceeded or, at least, approached to enable bonding. Such hot-press can produce initially a very satisfactory bond between the carbon cloth/paper and the ionomeric membrane. However, as the complete cell is assembled and humidified, the membrane swells significantly on rehumidification, whereas the carbon does not change dimensions on exposure to water. Therefore, the probability of partial delamination at the porous carbon/membrane contact on rehumidification is substantial, with consequences of loss of active surface area and/or increased cell resistance (recall that the liquid layer between a partly delaminated carbon electrode and membrane is deionized water).

8.3.5.1 Pt/C Catalysts Applied to the Membrane

In the early 1990s, an alternative low platinum loading catalyst layer structure was subsequently developed at the Los Alamos National Laboratory (LANL). In this structure, recast ionomer (rather than PEFE) is used to hold the catalyst layer structure together, and the low-loading catalyst layer is applied to the membrane

surface, rather than to the gas-diffusion structure (mode A3) [10–12]. Such (PTFE-free) layers have been described as ''thin-film'' catalyst layers because the high performance is obtained with a low catalyst loading (typically 0.2 mg Pt/cm^2 of electrode area) in a thin layer (<10 μm thick). By virtue of their thinness and the their optimized placement, high catalyst utilizations are obtained in these catalyst layers, and the continuity and integrity of the catalyst layer/membrane bond is improved compared to the structures prepared by modes B2 or B3. Although no hydrophobic component is introduced into this thin-film structure, gas permeability through it is sufficient to achieve good catalyst utilization at high current density in layers that are just 5–7 μm thick, thereby providing high cell performance [10–12].

The thin-film catalyst layers are typically cast from ''inks'' consisting of the supported Pt catalyst and solubilized ionomer [10]. Because the ionomer must bind the thin film structure together, special treatments of the recast films are necessary during fabrication to impart robustness to the ionomer in the catalyst layer. A thermoplastic form of the ionomer was proposed as an effective solution for the film robustness problem. Nafion® (and other perfluorinated ionomers) can be converted to a thermoplastic form by the ion-exchange inclusion of large, ''hydrophobic'' counterions such as tetrabutylammonium (TBA^+) [11]. The material is, consequently, moderately melt-processable. In an improved version of the original CCM fabrication process [11], thin-film catalyst layers are cast from an ink that consists of supported platinum catalyst and solubilized ionomer in the TBA^+ form. A general scheme describing the preparation of a catalyzed membrane employing thin-film catalyst layers is given in Fig. 22.

As seen in Fig. 22, MEAs based on such thin-film catalyst layers can be constructed using a ''decal'' process, in which the ink is cast onto Teflon blanks for transfer to the membrane by hot-pressing. A second approach is to cast the same type of ink (TBA^+ form of the ionomer) directly onto the membrane [12]. The latter process has an advantage over the decal process in

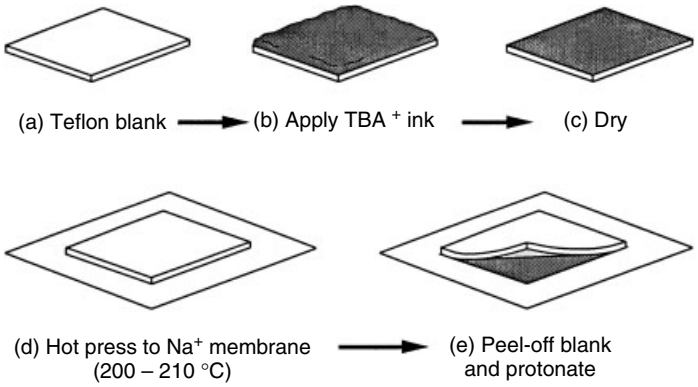


Fig. 22 Scheme describing preparation of catalyzed membranes for PEFCs by decal application of thin films (5–7 mm) of Pt/C//ionomer composite to the ionomeric membrane [11, 12].

the formation of a more intimate membrane/electrode interface. In the last step, the catalyzed membranes are rehydrated and ion-exchanged to the H^+ form by immersing them in lightly boiling 0.5 M sulfuric acid for several hours, followed by rinsing in deionized water. PEFC MEAs prepared by this technique have achieved the combined properties of very high catalyst utilization, and stable long-term performance (4000–5000 hours) with several shut-off/restarts and with full performance recovery after the cell has been exposed to several freeze/thaw cycles. PEFC polarization curves obtained with such thin-film catalyst layers applied by decal transfer to Nafion 117 and Nafion 105 (DuPont) membranes, as well as two other poly(PFSA) membranes, are shown in Fig. 8 for both H_2/air and for H_2/O_2 operation. The platinum loading of the cathodes for the four MEAs were 0.13, 0.14, 0.16, and 0.13 mg Pt/cm^2, respectively, and the anode catalyst loading somewhat lower. Also shown in these figures are the high-frequency resistances of the cells. Figure 8 shows the high PEFC performance achieved by the mid-1990s with low loadings of Pt applied by the thin-film, LANL technique, reaching current densities as high as 3 A cm^{-2} with oxygen cathodes and 1 A cm^{-2} at 0.6 V with air cathodes in PEFCs employing thinner ionomeric membranes. With highly purified H_2 as the anode feed stream, the total loss in cell power measured for PEFCs based on such catalyzed membranes after 3000 hours of continuous operation at 0.5 V was only 10% [51]. The gas reactants are reported (Fig. 8) to have been pressurized to 5 and 3 atm, respectively, thereby allowing higher cell performance. However, these testing conditions were

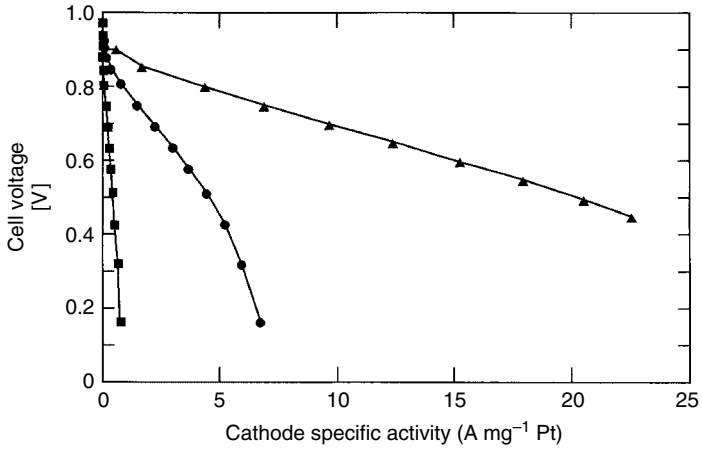


Fig. 23 Air cathode catalyst mass utilization (A mg^{-1} Pt) for different types of catalyst layers as developed chronologically for hydrogen/air PEFC. Squares: PTFE-bonded Pt black at 4 mg Pt/cm^2; circles: ionomer-impregnated, PA- type electrodes (0.45 mg Pt/cm^2); triangles: thin-film Pt/C//ionomer composite (0.13 mg Pt/cm^2). The relative advantage of thin-film catalyst layers is seen to increase with cell current density, as expected from the lower transport limitations involved (see Sect. 8.3.7.2.3) [10, 11].

set at the time without full consideration of the overall system optimization for maximum efficiency and minimum size and weight. As of today (a decade later), a similar performance at similar, or even lower, catalyst loadings has been reported with both gaseous reactants at ambient pressure, as discussed in the following. Comparison of catalyst utilization obtained with the different PEFC catalyzation techniques is given in Fig. 23, with all numbers applying to pressurized operation of the cells. The advantage in catalyst utilization with the thin-layer approach can be clearly seen, the utilization increasing with this approach particularly at the higher cell currents (lower cell voltage) owing to the minimized transport limitations in the thin catalyst layer.

Figure 24 shows a cross section of a Nafion membrane catalyzed by direct application of catalyst inks to its two major surfaces, as observed by SEM [52]. The thin slice of MEA required for SEM imaging was generated by microtome from the MEA encapsulated in epoxy. This figure actually describes an MEA prepared for a DMFC, with PtRu black and Pt black catalyst layers of relatively high loading, resulting in catalyst layers 10 and 14 μm thick (Fig. 24). The SEM image well depicts two generic characteristics of CCMs prepared by direct, ink-based application of the catalysts to the ionomeric membrane: the interface between the catalyst layer and the membrane is sharp on the SEM scale and the thickness of the catalyst layer measured from the

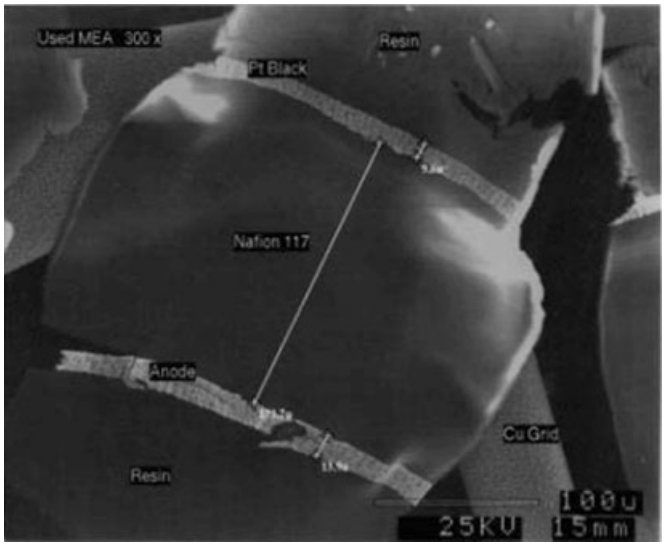


Fig. 24 A microtomed slice of a catalyzed Nafion™ 117 membrane used in a DMFC, imaged with an electron microscope. Two features gleaned from this image have been the sharpness of the catalyst layer/membrane interface (on the scale of the image shown) and a measured thickness of the catalyst layers about twice that calculated from perfect space filling by the composite catalyst layer material. The latter suggests a high fraction of free volume, facilitating gas transport [52].

image corresponds to a high-void volume fraction of 60–70% in the specific case shown.

The further, most recent significant advancement in the direction of thinner PEFC catalyst layers is based on a completely new concept of catalyst application to the polymeric membrane surface, demonstrated recently at the 3M company [53]. According to this new approach, an array of about 1-μm-long nanowhiskers, pointing up from the substrate surface, is first formed by vacuum deposition of a pigment (perylene red) onto a temporary substrate followed by thermal annealing. The Pt or Pt alloy catalyst is subsequently sputtered onto this supported array of nanowhiskers and the catalyst-coated whiskers are next pressure-embedded into the membrane surface, followed by removal of the decal support. Microimages of the nanowhisker array and of a cross section of the catalyzed membrane formed (in contact with an adjacent gas–diffusion layer) are shown in Figs 25 and 26, respectively. The limited length of the whiskers defines a very thin (about 1 μm thick) catalyst layer and, since the surface area (roughness) of the sputtered metal film itself around the whisker is relatively low, the total metal surface area in such nanostructured thin layers is, typically, 10–30 cm^2 Pt/cm^2 of membrane cross-sectional area – an order of magnitude lower versus that typically achieved with 2-nm catalyst particles applied to the membrane as composite Pt/C//ionomer layers. However, it turns out that the *specific* electrocatalytic activity (mA cm^{-2} Pt) exhibited in the cathode process (oxygen reduction) by a relatively smooth Pt film sputter-coated onto such nanowhiskers is measured to be five to six times higher than that of 2 nm, pseudospherical Pt particles used to generate larger Pt surface areas. As a result, exceptional activities per milligram of Pt have been demonstrated with the 3M mode of CCM fabrication at ultralow total catalyst loadings of 0.1 mg Pt/cm^2 of membrane area. High PEFC

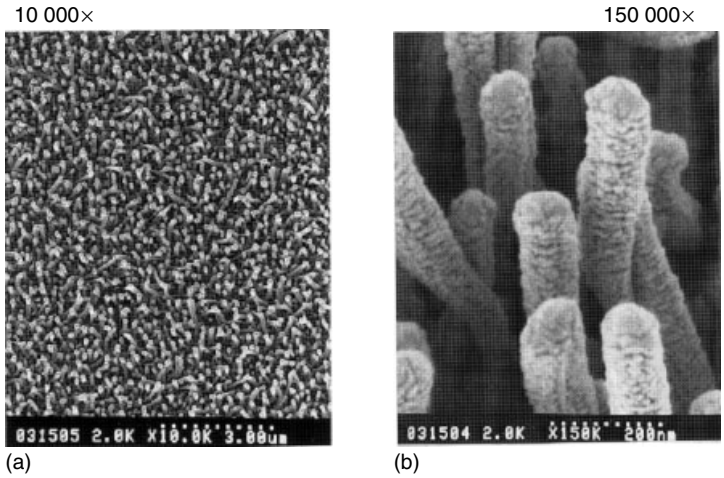


Fig. 25 Nanostructured ultrathin film catalyst developed at 3M. (a) Micrograph at 10 000 magnification illustrating the uniaxial orientation and the high number density of Pt-coated whiskers and (b) 150 000 magnification view at 45° viewing angle of the whiskers in (a) [53a].

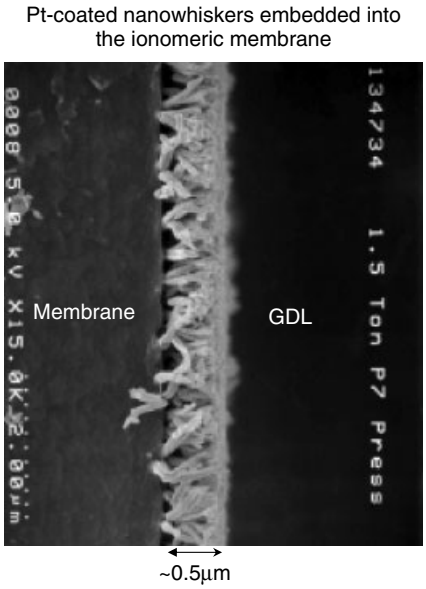


Fig. 26 Pt-coated nanowhiskers (Fig. 25) embedded into the ionomeric membrane and backed by a GDL to form an MEA with an ultrathin catalyst layer [53b].

performance on ambient pressure air and hydrogen at total cathode + anode loading of just 0.12 mg Pt/cm^2, achieved with such nanostructured thin-film catalyst layers, is demonstrated in Fig. 27. The significantly higher specific catalytic activity (mA cm^{-2} Pt at given V_{cath}) of such relatively smooth Pt films formed around nanowhiskers can be understood in light of other reported observations of lowered activity per catalyst surface area exhibited by Pt particles as small as 2 nm in diameter. The apparent reason for the "diminishing returns" seen on high dispersion of Pt catalysts beyond the level of 20–50 m^2 g^{-1} Pt are nonoptimized strengths of surface-reaction intermediate bonds to a relatively large population of edge sites on such small metal particles. In contrast, the morphology of the Pt catalyst film around the nanowhisker indeed corresponds to a lower metal surface area but thereby avoids particle size–related drawbacks of very small individual catalyst particles. It can be readily understood (see Fig. 4) that the high performance of such ultrathin

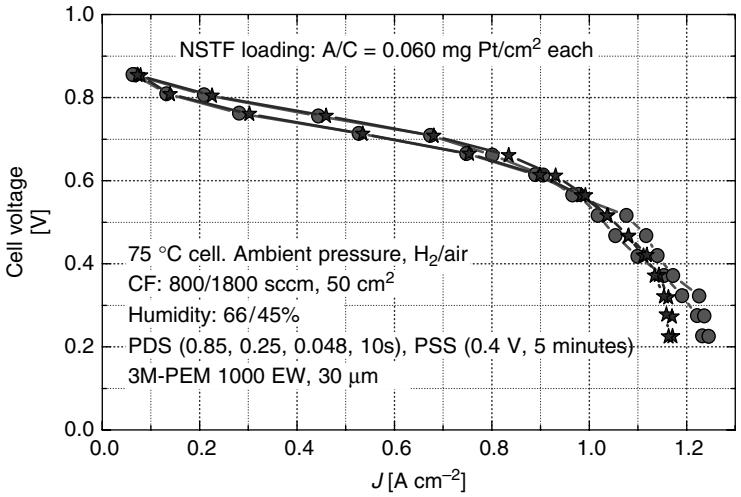


Fig. 27 Hydrogen/air PEFC performance at ambient pressure of both fuel and air, achieved at total Pt loading of only 0.12 mg Pt/cm^2 with an MEA based on an ultrathin catalyst layer of the type described in Figs 24 and 25 [53b].

catalyst layer is strongly assisted by further shortening of the protonic, electronic, and reactant transport paths to all catalyst sites in the 1-μm-thick layer. However, limitation of such strongly reduced overall mass of catalyst arises in cases of fuel-cell electrocatalytic processes slower than those in hydrogen/air cells at a cell temperature over 70 °C. And, another obvious concern is the ever-growing sensitivity to some given level of impurity when the catalyst surface area is dropped to such low levels. On the other hand, an extra advantage of this type of catalyst is that the pigment support employed seems to be much less susceptible to electrooxidation than carbon (see section 8.3.5.2), consequently existing significantly better stability on the cathode side of the PEFC.

8.3.5.2 Long-term Stability of Membrane/Electrode Assemblies in PEFCs

A fuel-cell stack developed for transportation applications has to maintain a stable performance for thousands of hours while operating on realistic fuel feed streams, that is, feed streams possibly containing significant levels of impurities. And, the long-term performance stability demand on fuel-cell systems for stationary power generation (distributed generation) is an order of magnitude more severe, that is, of the order of 10^4 hours. Critical questions raised in this regard are as follows:

1. Will the electrocatalytic activity be maintained in spite of the potential for (a) catalyst loss and/or catalyst agglomeration phenomena and (b) catalyst surface poisoning?
2. Is the integrity of the ionomeric membrane (specifically the poly(PFSA) membrane) within an operating PEFC, sufficient to prevent damage, for example, pinholes, during prolonged operation (with such damage possibly leading to hazardous, direct fuel/oxidant intermixing)?
3. Can the high protonic conductivity of the membrane be maintained for thousands of hours in spite of any possible source of trace ionic contaminants that could replace protons in the ionomer?

Reasonably satisfactory, albeit by no means complete, answers have been provided for the above questions. Life testing of pressurized PEFC stacks developed for space application at GE (later Hamilton Standard) first demonstrated thousands of hours of operation without failure [54]. A more recent, 11 000 hours test of a PEFC eight-cell stack with neat oxygen and neat hydrogen feeds showed a low decay rate of 1.3–1.4 μV $hour^{-1}$ under constant load [55]. In the last two cases, the catalyst loadings employed were high: at least 4 mg of Pt black per square centimeter on each side of the cell. The decay rate with order of magnitude lower catalyst loadings, as required for cost competitiveness in most PEFC applications, is still to be fully quantified for specific stack materials and builds and under various operation conditions. Oxidation of the carbon support material used in the Pt/C catalysts, typically employed in PEFCs of lower loadings, is one cause of faster deterioration in cells employing such lower loadings [56]. Highest rates of cathode carbon oxidation are measured at or near the open-circuit potential and are strongly enhanced at an elevated cell temperature. Under open-circuit conditions, the cathode is not protected by the lower electrode potential established under current (“providing cathodic protection”) and the Pt-catalyzed, carbon-support oxidation rate then becomes highest, with

possible results ranging from loss of cathode catalyst layer transport characteristics to overall catalyst layer disintegration. Possible countermeasures are the use of more graphitized carbon supports that are less vulnerable to oxidation and/or optimization of the operation conditions, for example, leaving the cell under load until it cools down significantly to provide maximum cathodic protection to the cathode as long as the cell temperature is elevated.

The Pt catalyst itself is also found to be vulnerable to some loss by corrosion, involving oxidation of Pt metal atoms from the catalyst particle to generate Pt ions, that is, Pt particle dissolution. Recent work [57, 58] has revealed a behavior of Pt in this regard that is not unlike that of more reactive metals: Pt can be, within the potential range relevant for PEFC or DMFC electrodes, in either a ''passive state'' or in a more ''active state'' with regard to metal dissolution rates. A passive state of the Pt catalyst in a PEFC, cathode, is established on steady operation at potentials above 0.8 V versus the hydrogen electrode, at which potential a protective superficial surface oxide layer forms on the Pt catalyst surface, reducing the rate of metal dissolution to very low levels. The ''active state'' of Pt occurs, on the other hand, as result of the reduction in this surface oxide layer, typically by operation for some period of time at a lower electrode potential, for example, 0.5 V versus hydrogen. Once a Pt particle surface is ''activated'' in this way, fast subsequent perturbation of the cathode potential back to values higher than 0.8 V results in unusually high instantaneous rate of Pt dissolution, lasting until the passive state is reestablished by surface oxide formation.

Pt catalyst ripening in PEFCs electrodes, resulting in loss of catalyst active area, was described earlier on [51]. Catalyst ripening was measured in PEFC cathodes of low Pt loading, that operated continuously for 2500 hours on pressurized hydrogen and air at a cell voltage of 0.5 V [51]. Using TEM imaging, measurements taken for the cathode catalyst showed that slow catalyst ripening took place, resulting in as a decrease in Pt surface area from an initial value of 100 $m^2\ g^{-1}$ to 70 $m^2\ g^{-1}$ after 1000 hours and to 40 $m^2\ g^{-1}$ after 2500 hours. The results of particle size distribution analyses for as-supplied Pt/C catalysts and for such catalysts after prolonged use in the PEFC cathode are shown in Fig. 28. Such surface area loss (Fig. 28) suggests that the cathode catalytic activity should have declined by about 60% between the beginning and the end of the life test. However, long-term losses of the cell current measured at 0.8 V between beginning and end of the life test (2500 hours) were practically negligible. A likely explanation is that such loss of Pt surface area is compensated by a higher specific activity per unit Pt surface area (mA cm^{-2} Pt) of the larger Pt catalyst particles formed in the particle agglomeration process.

The vulnerability of Pt and Pt alloy catalysts to poisoning by trace contaminants at operation temperatures typical for a PEFC is well documented and is of clear concern in the design of a power system based on a PEFC stack. Sources of contaminants include both fuel and air feed streams as well as processes derived from chemical instability of cell component(s). As to the feed streams, ''polishing'' of anode feed streams generated by fuel processing upstream the cell should leave very low levels of CO to be dealt with effectively within the cell (see Sect. 8.3.7.1), whereas any traces of sulfur or ammonia have to be perfectly eliminated upstream the anode

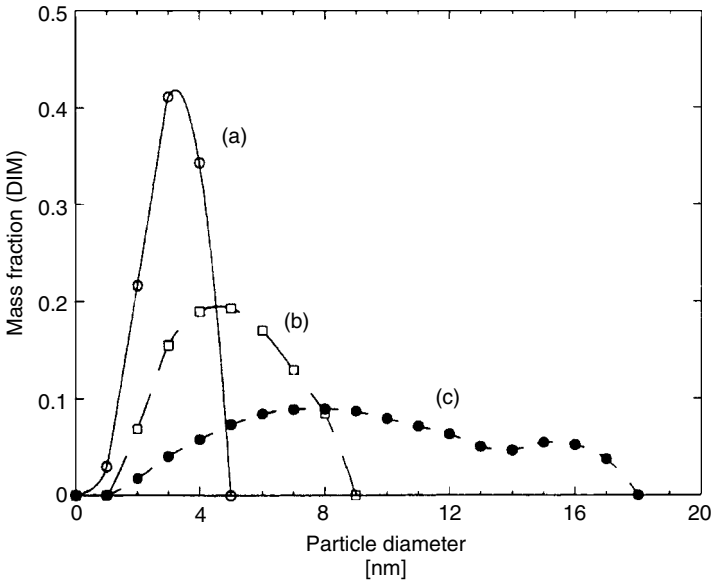


Fig. 28 Mass distributions of carbon-supported Pt catalyst particles (a) in the as-cast catalyst layer, (b) after 1300 hours of operation, and (c) after 2200 hours of operation as PEFC cathode catalyst [51].

[59], which requires, in the case of some fuel sources, the introduction of dedicated traps/filters. Filters are also routinely used for the air feed stream to minimize entry of any hydrocarbon, sulfur, nitrogen oxides, and other "bad players" into the cathode, thereby readily causing catalyst surface poisoning. The common approach to minimize any question marks on possible catalyst poisoning originating from cell hardware components, is to prefer graphite or composite carbon/polymer structural elements to metal elements. The latter can be a source of trace metal ions that could poison the Pt or Pt-based catalyst, particularly in cases with a strong tendency for "underpotential deposition". However, superior mechanical properties of metal components allowing to achieve structural integrity with thinner profiles provide a sufficient drive for their use in PEFC and DMFC stacks of higher power densities in spite of remaining corrosion concerns. To protect both the membrane and catalyst from possible detrimental effects of leached metal ions, requires great care in the selection of the structural materials and involves, in many cases, surface coating to achieve a well-protected metal surface. Some basic rules typically followed are (1) use of "more noble" forms of stainless steel for any stack or "plumbing" component of a PEFC system, for example, use of 316 L stainless steel as minimum (and avoidance of any less expensive alloys like 304) and (2) for stack structural elements and current collectors, application of a surface coat of good adhesion to the substrate metal surface that exhibits low contact resistance when pressed against the carbon-based backing layer of the MEA. Such possible metal surface coats include carbon-based coats, electronically conducting oxide-based coats,

or thin gold coats where cost is of lower concern.

The mechanical integrity of ionomeric membranes employed in PEFCs is challenged when the membrane thickness is taken down to 25 μm or even lower, as required for maximizing power output and facilitating water management. Reported life tests suggest that such thin poly(PFSA) membranes have sufficient intrinsic integrity, which may be improved further through the use of reinforcement by some open mesh, as fabricated, for example, by W.L. Gore. However, any imperfections in stack design and build, together with various possible modes of chemical attack on the membrane in the operating PEFC, result in membrane operational integrity that can fall significantly short of expectation based on mechanical properties of the dry, or the wet membrane. Furthermore, membrane failure will likely occur if PEFC stack design does not ensure sufficient humidification over the whole area of the membrane under the complete range of conditions. Local thermal runaway cannot be ruled out if local dehydration brings about increased local membrane resistivity and, in turn, increased heat generation next to highly resistive fraction(s) of membrane area. Membrane integrity can also be challenged by the (lateral) dimensional changes of membrane segments clamped between hardware elements of the stack, as result of ''dry-wet'' cycles. Counter measures include avoidance of severe dry-out at elevated temperatures and improvement of the hydro-rheological properties of the ionomeric membrane.

Membrane conductivity losses by ion exchange seem to be easier to prevent: only small losses of protonic conductivity, of the order of 5–10% after 4000 hours, have been observed in well-humidified cells during PEFC life tests according to measurements of cell impedance at 5 kHz [11]. However, the deionized water employed in the humidification scheme in such tests [11] had very low levels of metal ions (e.g., $Fe^{+2/+3}$, Ca^{+2}, or Mg^{+2}). Such multivalent ions exchange irreversibly with protons in the poly(PFSA) membrane, causing a drop in membrane conductivity. Deionizing the water used for PEFC humidification is, therefore, required, and appropriate plumbing materials should be used to avoid generation of ionic contaminants by corrosion processes.

A critical issue of chemical instability of the ionomeric membranes most commonly used in PEFC technology, repeatedly brought up recently, is the attack of poly(PFSA) membranes by $OH^{\bullet}$ radicals, followed by further attack by hydrogen peroxide [60]. There is good evidence for the formation of $OH^{\bullet}$ radicals in the PEFC cathode by recombination of hydrogen crossing over the membrane with oxygen at the cathode or vice versa with formation of the $OH^{\bullet}$ radical at the anode. Although the crossover rate of hydrogen and oxygen through the membrane is equivalent of only several milliamperes per square centimeter, it is apparently sufficient to cause enough initial damage to the structure of poly(PFSA) materials currently manufactured, and this initial attack by the $OH^{\cdot}$ radical apparently triggers further attack of the membrane by hydrogen peroxide. The latter is a not uncommon side product of the cathode process, particularly when the cathode catalyst surface is partly deactivated. Remedial efforts pursued at this time, inclue minimizing the probability of OH radical formation in the cell by maintenance of high membrane hydration through the thickness dimension (hence use of thinner membranes) as

well as protecting vulnerable sites in state-of-the-art poly(PFSA) polymeric material, suspected to be associated with polymer chain termini. Very recent reports suggest that cerium, or manganese ion additives provide protection against of radial attack in poly[PFSA] membranes.

Finally, the ability to withstand freeze/thaw cycles is another important aspect of the long-term performance stability demanded of an MEA in a PEFC or DMFC in any application involving stack exposure to "the elements". Recent reports from UTC Fuel Cells of PEFC stack recovery from deep freeze by just electrically loading the stack, that is, allowing self generated current to pass through the cells of the stack, have been of great interest. Those reports, also highlighted, however, fluctuations of ice crystals location and morphology within the stack during such freeze/thaw cycles, which could, in principle, cause some structural damage. MEA and overall stack vulnerability to freeze/thaw cycles is certainly a sensitive function of stack design, particularly regarding water management under load as well as when idling. The final word on the extent of any damage per given design, following more than just several cycles as demonstrated to date, is still awaited.

8.3.6 From a PEFC Stack to the PEFC Power System

Figures 3–5 describe how, by fabricating the framed MEA and stacking multiple MEA/bipolar plate building blocks containing flow fields for air and fuel streams (Fig. 2), the PEFC stack fabrication is completed. The stack is indeed the heart of any PEFC power system and the quality of a PEFC power system depends to large degree on the quality of the stack, including power output per unit weight coupled with a desirable conversion efficiency, long-term performance stability, and the probability of failure.

A fuel-cell *power system* requires however, additional component and, in some typical cases, such auxiliaries constitute the larger fraction of the power system. The simplest fuel cell systems are associated with direct conversion fuel cells and, in the case of the PEFC, such simple systems are based on direct conversion of 100% hydrogen fuel. The most relevant potential application is in transportation, where 100% hydrogen fuel is to be carried on board and fed directly to the anode side of the stack. As can be seen in Fig. 29a [61], the system components required in addition to the stack itself have functions of air and of water management. Air has to be passed continuously through the cathode side of the stack and this requires a fan, air pump, or an air compressor, depending on the stack design and on the required stack performance. Recent stack designs typically prefer to minimize air pressure to 1.5 atm (abs) as a compromise between the need to enhance the partial pressure of oxygen in air/saturated water vapor mixtures at cell temperatures of 70 °C or above and the counter requirement to minimize compressor losses.

Water management can be most simply achieved by providing the gas feed humidification level required to maintain the conductivity of the fuel-cell membrane and of the ionomers in the catalyst layers. Gas feed humidification has been achieved by a variety of methods including, for example, "enthalpy exchangers" [6] and porous bipolar plates [62]. The two latter approaches rely on utilization of stack-produced water, thereby eliminating the need of frequent "water refill". The system in Fig. 29a, can use a condenser to

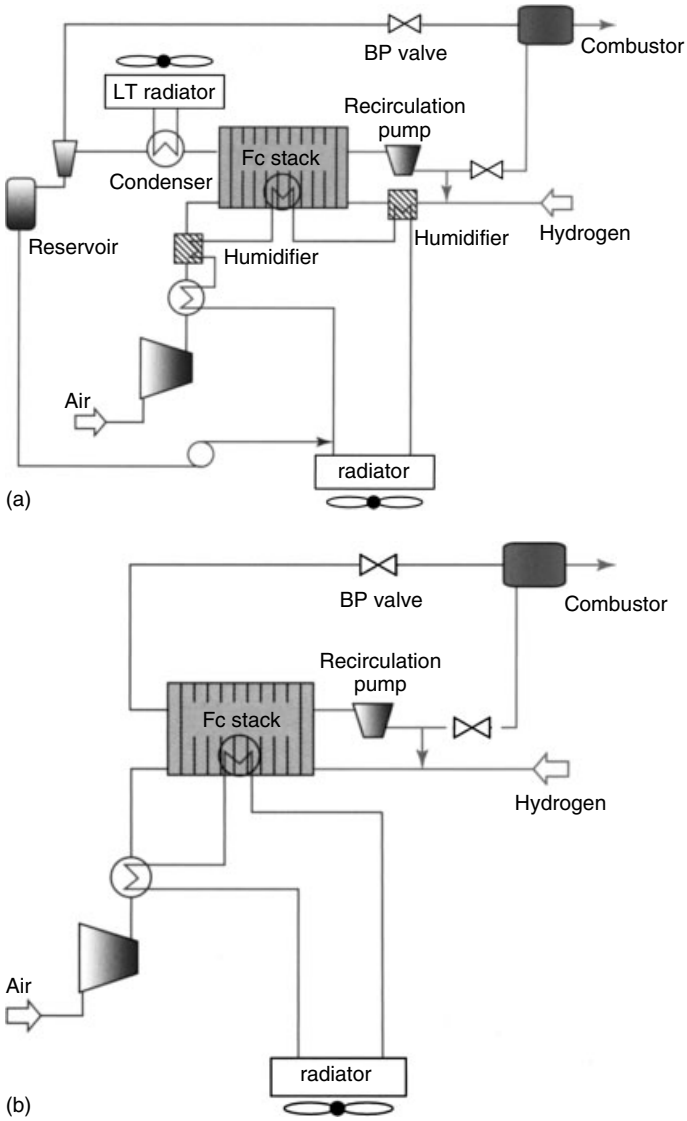


Fig. 29 Schemes of power systems based on a neat hydrogen/air, PEFC stack, for the case (a) where pre-humidification of the incoming fuel feed stream is required and (b) where satisfactory stack hydration is achievable w/o pre-humidification of the incoming fuel feed stream [61].

ensure sufficient capture of water from the cathode exhaust so as to achieve water mass balance with no need of water supply from the outside. The liquid water is then collected and pumped externally to the cell into two humidifiers for the incoming air and the incoming hydrogen streams, respectively, thereby ensuring effective humidification across the active area of each cell. The same function of

water exchange between exhaust and inlet streams can be basically achieved within the stack, using concepts like the ones described in Ref. 6 or 62. When this is effectively implemented, there is no need to humidify the inlet feed streams, either fuel or air, in front of the stack. The system becomes, as result, significantly simpler, as showin in Fig. 29b [61].

In contrast to such relatively simple systems shown in Figs. 29a, b a PEFC power system using some ''hydrogen precursor'', for example, natural gas, which needs to be processed upstream the PEFC to generate a hydrogen-rich gas mixture as feed stream to the stack, is significantly more complex [63]. A generic block diagram for such a power system is provided in Fig. 30. Multiple-step processing of the fuel at high temperatures, including reforming, shift reaction, and selective oxidation, is required to convert a hydrocarbon fuel to the required mixture of hydrogen, carbon dioxide, nitrogen, and water vapor fed to the PEFC stack. The complexity of fuel processing has a lot to do with the need to lower the level of carbon monoxide in the processor outlet stream to accommodate a fuel cell operating under 100 °C (see Sect. 8.3.7.1). The obvious rationale behind such a complex system is the ready availability of, for example, natural gas as fuel for a power system, whereas the 100% hydrogen fuel that could enable a much more simple system is not readily available at this point for any large-scale application.

In a PEFC power system, losses in fuel processing upstream the cell add to the losses in the stack itself and the result is that, when starting from a fuel like

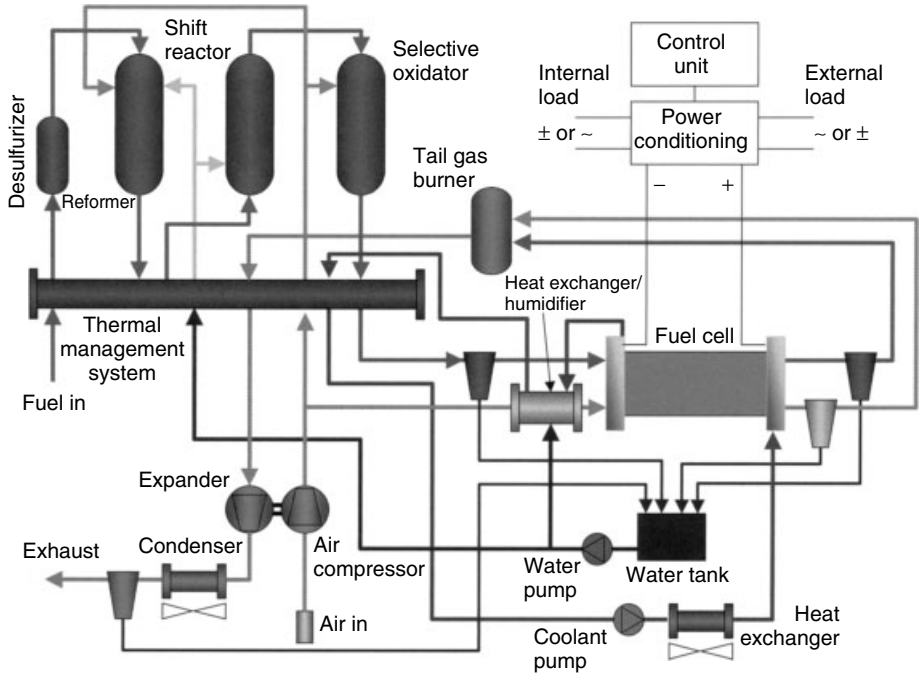


Fig. 30 Scheme of a power system based on a PEFC stack operating on reformed carbonaceous fuel (eg., natural gas) and air [63].

natural gas, the overall efficiency of the system is not better than 70–80% of the efficiency of the stack, that is, with a PEFC stack operating, for example, at 60% conversion efficiency, the overall conversion efficiency of natural gas-to-hydrogen–rich gas mixture–to-electric power is not likely to be better than 42–48%. Successful implementation of a PEFC power system employing a tank of carbonaceous fuel, becomes to a large degree a function of the fuel processing subsystem, and the pretty complex fuel, air, water, and heat management required in a system operating between a temperature close to 1000 °C (natural gas reformer) and 70 °C (PEFC stack). Further detail about fuel-cell systems can be found in Refs 3 and 5.

8.3.7 Electrocatalysis and Electrodes in PEFCs

8.3.7.1 The Anode Process in the PEFCs

The rate of the hydrogen oxidation process at the Pt/ionomer interface at 80 °C is very high (a current density of 1 A cm^{-2} is obtained at a anode overpotential of a few mV) as long as the catalyst surface is not contaminated by adsorbed impurities. Nevertheless, in actual PEFC systems, the anode typically presents significant challenges regarding the very high sensitivity of Pt and Pt alloy anode catalysts to fuel impurities at the cell temperatures typical for PEFCs. It is particularly the strong sensitivity of the anode catalyst to traces of CO in the fuel feed stream, that makes ''good CO tolerance'' a property of high premium in PEFC technology, particularly when the hydrogen feed is derived by processing of carbonaceous fuels. The maximum power output of a PEFC drops by almost an order of magnitude when the feed is changed from pure hydrogen to a diluted hydrogen feed stream containing 100 ppm CO. Such high sensitivity to this very common impurity highlights a disadvantage of this low-temperature fuel-cell technology that accompanies all the important advantages rendered by low temperature of operation.

The science and technology of the PEFC anode have consequently focused on the issue of impurity tolerance, targeting high-performance levels in PEFCs operating on processed carbonaceous fuels. Processed natural gas (methane) has always been considered the appropriate feed stream in stationary power applications (''distributed generation'') of PEFCs and, particularly during the 1990s, not only methanol but even gasoline was considered as possible feed for a fuel processor on board a vehicle generating hydrogen-rich gaseous mixtures for PEFC vehicular power sources [64]. In PEFC systems based on autothermal reforming of gasoline, or of methane, it would be required to achieve high cell and stack performance with anode feed streams as low as 40% in hydrogen and containing CO and other critical impurities, such as ammonia, at levels that will strongly affect performance and performance stability, unless dedicated, additional system components and processes are introduced.

8.3.7.1.1 The PEFC Anode with Pure Hydrogen Feed

Probably as a result of the high performance of (pure) hydrogen Pt anodes in PEFCs, relatively little work has been done on the rate and mechanism of the hydrogen oxidation process at an impurity-free Pt/ionomer interface. Impedance spectroscopy work performed by Raistrick [65] on a model system of a flat Pt electrode pressed against a Nafion membrane, revealed details of the mechanism of hydrogen oxidation at the Pt/ionomer interface at room temperature.

An example of an impedance spectrum measured by Raistrick for the H_2/H^+ couple at the smooth Pt electrode/hydrated Nafion™ membrane interface at room temperature is given in Fig. 31. The spectrum includes three features which correspond, in order of increasing frequency, to (1) mass-transport-limited supply of H_2 gas to the planar Pt electrode pressed against the membrane, (2) the dissociative chemisorption of H_2 molecules on Pt catalyst sites to form adsorbed hydrogen atoms, and (3) electrooxidation of adsorbed hydrogen atoms. The resolution of the latter two steps in the overall interfacial process of hydrogen electrooxidation can be contrasted with the impedance spectrum for the ORR the smooth Pt/Nafion membrane interface (see Fig. 47), which includes only a single feature associated with interfacial charge-transfer indicating a single rate-limiting interfacial charge-transfer step. The interfacial process of H_2 electrooxidation seems to be controlled at higher current densities by the rate of the dissociative chemisorption step [65], and this impedance spectral feature is also most strongly affected by the introduction of traces of carbon monoxide into the hydrogen gas [65]. The total impedance associated with the interfacial hydrogen oxidation process at the Pt/hydrated Nafion™ membrane interface is much smaller compared to that of the ORR process in the same electrolyte. The total faradaic impedance at the equilibrium potential for the H_2/H^+ couple is seen in Fig. 31 to be around 25 $\Omega\ cm^2$, corresponding to an exchange current density of the order of 1 mA cm^{-2} Pt for the H_2/H^+ couple at the Pt/Nafion membrane interface at room temperature. A similar net rate for the ORR at this interface is achieved only at cathodic overpotentials exceeding 0.4 V, which means that the electrocatalytic ORR process is about 4 orders of magnitudes slower.

As a result of the high rate of the hydrogen oxidation process at Pt in contact with this ionomeric medium, the interfacial potential drop at a well-humidified H_2 anode in a PEFC operating

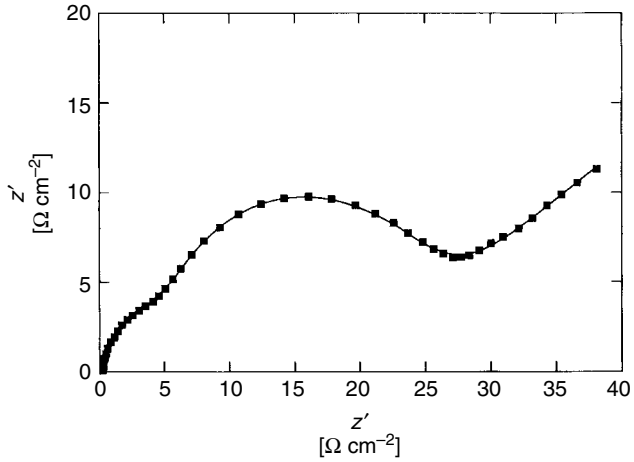


Fig. 31 Impedance spectrum measured for a bulk, smooth Pt/Nafion membrane interface in a hydrogen atmosphere at the H_2/H^+ equilibrium potential [65].

at 80 °C is very low: for a hydrogen exchange current density of 10 mA cm^{-2} at 80 °C and with 400 cm^2 Pt/cm^2 of cross-sectional area of the MEA (from, e.g., 0.3 mg of (2 nm) Pt particles per square centimeter of cross-sectional area), the effective anode exchange current density is 4 A cm^{-2} and the anode overpotential at cell current of 1 A cm^{-2} can be calculated from the linear current density versus overpotential expression as

$$\eta_{\text{anode}} = (RT/F)J/J_0 \qquad (23)$$

yielding η_{anode} of just 6 mV at what is typically the PEFC current density at peak power.

However, the case will be quite different if (1) the anode side of the membrane becomes dehydrated (see Sect. 8.3.4.4.3) and/or (2) the hydrogen feed stream to the cell is less than perfectly pure. The latter point is discussed in detail next.

8.3.7.1.2 **The Anode Process in a PEFC Operating on Reformed Carbonaceous Fuels**

PEFC power systems for stationary power generation, as well as some power systems for transportation, have relied on processing of liquid fuels on board the vehicle to generate a hydrogen-rich gas mixture as the feed stream for a PEFC stack. Steam reforming of methanol and autothermal reforming of gasoline-like fuels have been the two most commonly considered and investigated systems of fuel processing on board vehicles, whereas methane processing has been developed for stationary applications. Methanol has the advantage of a relatively low-temperature steam-reforming process, yielding a gas mixture with hydrogen at levels as high as 75% of dry gas composition. Gasoline, on the other hand, was looked at for several years as a possible source of hydrogen, although it requires significantly higher temperatures for autothermal processing and the latter process yields typical compositions of 35–40% hydrogen in the dry gas mixture. The obvious incentive to tackle such (very) high technical hurdles in the development of a gasoline on board processing subsystem was the hope to rely on the present infrastructure of fuel for transportation. In both of these fuel-processing systems, water shift and oxidative fuel cleanup steps follow the steam or autothermal reforming step. These steps lower CO levels in the product gas mixture from several percents down to the 10–100 ppm level, as required for use as anode feed stream in a PEFC operating at 80 °C. The anode feed stream in such systems is thus typified by hydrogen mole fractions (before humidification) ranging between 0.4 and 0.75 and by the presence of 10–100 ppm CO, a serious poison of Pt-based anode catalysts operating at 80 °C.

The effect on PEFC performance of 10–100 ppm CO in the anode feed stream is very substantial when Pt catalyst is used in the anode. It can be seen from Fig. 32 that the effect of CO can be qualitatively described in terms of some ''critical current density'' of value which drops with the level of CO in the anode feed stream. When current demand is below the critical current level, the PEFC maintains practically CO-free cell performance, whereas above that critical current level, cell performance drops sharply as the cell voltage falls well under the corresponding CO-free level. The explanation for this behavior was provided by a model that considered the interfacial kinetics at the PEFC anode in the presence of low levels of CO [42, 66, 67].

Four processes, expressed in the following Eqs (1–4), were assumed to determine the interfacial kinetics [66]. By itself, M

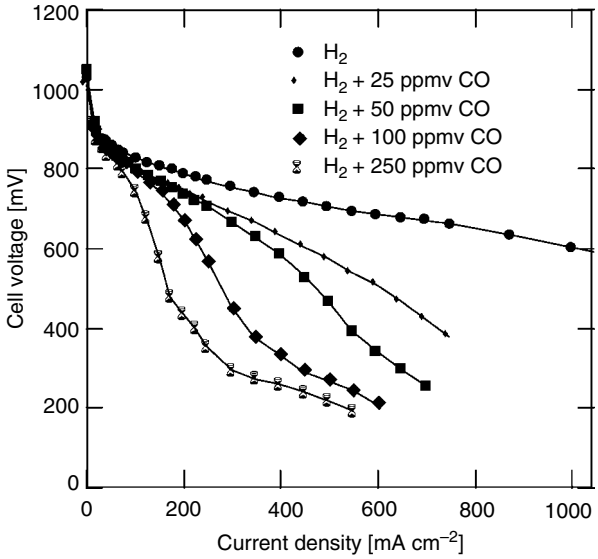


Fig. 32 Measured, iR-corrected, hydrogen/air PEFC voltage versus current density in operation at 80 °C on high stoichiometric flow of hydrogen with a range of concentrations of carbon monoxide. Pt catalyst was at 0.2 mg cm^{-2} on each side [68].

corresponds to a vacant catalyst site, otherwise either H or CO is attached to the site.

$$\mathrm{CO} + \mathrm{M} \underset{b_{\mathrm{fc}}(\theta_{\mathrm{CO}})k_{\mathrm{fc}}}{\overset{k_{\mathrm{fc}}}{\rightleftharpoons}} (\mathrm{M}-\mathrm{CO}) \qquad (24)$$

$$\mathrm{H_2} + 2\mathrm{M} \underset{b_{\mathrm{fh}}k_{\mathrm{fh}}(\theta_{\mathrm{CO}})}{\overset{k_{\mathrm{fh}}(\theta_{\mathrm{CO}})}{\rightleftharpoons}} 2(\mathrm{M}-\mathrm{H}) \qquad (25)$$

$$(\mathrm{M}-\mathrm{H}) \xrightarrow{k_{\mathrm{eh}}} \mathrm{H^+} + e^- + \mathrm{M} \qquad (26)$$

$$\mathrm{H_2O} + (\mathrm{M}-\mathrm{CO}) \xrightarrow{k_{\mathrm{ec}}} \mathrm{M} + \mathrm{CO_2} + 2\mathrm{H^+} + 2e^- \qquad (27)$$

The first two equations represent the competing processes of CO adsorption and dissociative chemisorption of H_2. As seen, it was assumed [66] that desorption rate for CO and the adsorption rate for hydrogen could both be functions of θ_{CO}. The third process (26) generates the current corresponding to anodic hydrogen oxidation, that is, electrochemical oxidation of adsorbed hydrogen atoms. The fourth process (27), corresponding to the electrochemical oxidation of CO to CO_2, reaches measurable rates only at higher anode overpotential. Nevertheless, very low rates (e.g., microamperes per square centimeter of MEA cross-sectional area) of process (27) expected at anode potentials well under the value normally considered as ''onset'' or ''threshold'' of CO electrooxidation, are calculated to be sufficient to significantly increase the rate of process (25) and, consequently, the overall rate of hydrogen oxidation in the presence of low levels of CO. The anode overpotential required to accelerate process (27) to some significant

rate would depend primarily on the potential (versus RHE) corresponding to a similar rate of "water activation" to form active oxygen intermediate on the surface of the anode electrocatalyst. Anodically formed surface oxygen will act subsequently readily with adsorbed CO to complete process (27). Step (27) is indeed commonly described as a sum of two substeps:

$$M' + H_2O \longrightarrow M' - OH + H^+ + e^- \quad (28)$$

$$M - (CO) + M' - OH \longrightarrow CO_2 + H^+ + e^- + M' + M \quad (29)$$

On a Pt catalyst particle, M′ and M are both (obviously) Pt surface sites, although CO adsorption and water activation may take place on distinguishable types of sites for example, CO adsorption on crystal surface terraces and OH formation on steps, or edges, with the process then being completed by migration of the mobile, adsorbed CO molecules along the terrace to the step to recombine with the surface OH groups. In the case of PtRu alloys, known to enable onset of process (28) at significantly lower anode potentials versus unalloyed Pt, M′ in Eq. (28) is likely a Ru site, whereas M is likely an adjacent Pt surface site.

On the basis of the scheme described by Eq (24), Eqs. (30) and (31) represent the steady-state catalyst surface balance of adsorption, desorption, and electrooxidation fluxes of carbon monoxide and of hydrogen. If the intermediate hydrogen step is second order in catalyst sites, as was assumed on the basis of the nature of process (25) [66], then, in Eq. (31), $n = 2$. Equations (30) and (31) determine θ_{CO}, the fraction of catalyst sites with adsorbed CO, and θ_h, the steady-state fractional coverage with adsorbed H at steady state, at anode potential V_a. The fluxes are also all in units of equivalent current density. The quantity ρ represents the molar area density of catalyst sites times the Faraday constant.

$$\rho\dot{\theta}_{CO} = k_{fc}x_{CO}P_A(1 - \theta_{CO} - \theta_h) - b_{fc}k_{fc}\theta_{CO} - k_{ec}\theta_{CO}e^{\frac{V_a+V_{eq}}{b_c}} = 0 \quad (30)$$

$$\rho\dot{\theta}_h = k_{fh}x_hP_A(1-\theta_{CO}-\theta_h)^n - b_{fh}k_{fh}\theta_h^n - 2\theta_h k_{eh}\sinh\left(\frac{V_a + V_{eq}}{b_h}\right) = 0 \quad (31)$$

$$j_h = 2k_{eh}\theta_h\sinh\left(\frac{V_a + V_{eq}}{b_h}\right), \quad j_{CO} = 2k_{ec}\theta_{CO}e^{\frac{V_a+V_{eq}}{b_c}} \quad (32)$$

Current density expressions for hydrogen and CO oxidation, j_h and j_{CO}, are given in Eq. (32) as a function of anode potential V_a referenced to a SHE corrected for local Nernstian potential shift associated with local values of p_{H_2} different than 1. Note that the same overpotential is assumed at a given anodic potential for both H_{ads} and CO_{ads} electrooxidation. This means that both k_{fh} and k_{fc} describe the rate of the anodic process at 0 V RHE (and full surface coverage by the corresponding surface species).

If all the rate constants are independent of θ_{CO}, that is, k_{fh} and b_{fc} are constant, then an analytic solution for θ_h and θ_{CO} is given by Eq. (33) with intermediate variables defined in (34) [66]. From the expressions for j_h and j_{CO} in Eq. (32), the kinetic current densities of hydrogen and of CO electrooxidation can be evaluated.

$$\theta_H = \frac{\left(a_1\sqrt{a_2a_3k_{fh}P_Ax_H + a_1^2s_1^2}\right) - a_1^2s_1 - a_2k_{fh}P_Ax_H}{k_{fh}\left(b_{fh}a_1^2 - a_2P_Ax_H\right)}$$

$$\theta_{CO} = \frac{k_{fc}P_Ax_{CO}\left(a_1b_{fh}k_{fh} + a_1s_1 - \sqrt{a_2b_{fh}k_{fh}^2P_Ax_H + 2a_2k_{fh}P_Ax_Hs_1 + a_1^2s_1^2}\right)}{k_{fh}\left(b_{fh}a_1^2 - a_2P_Ax_H\right)} \quad (33)$$

$$s_1 = k_{eh}\sinh\left(\frac{V_a + V\text{eq}}{b_h}\right),$$

$$a_1 = e^{\frac{V_a}{b_c}}k_{ec} + k_{fc}(b_{fc} + P_Ax_{CO})$$

$$a_2 = \left(e^{\frac{V_a}{b_c}}k_{ec} + b_{fc}k_{fc}\right)^2,$$

$$a_3 = b_{fh}k_{fh} + 2s_1 \quad (34)$$

This analytical solution revealed well the qualitative ''critical current'' behavior seen experimentally (Fig. 32), but did not fully predict the quantitative dependence of the critical current at a Pt anode catalyst on P_{CO} in the anode feed [66]. The latter required assuming a Temkin adsorption isotherm for CO at the anode catalyst, as originally suggested in Ref. 67. By using a Temkin isotherm, Eq. (36) allows the free energy of CO adsorption to decrease linearly with θ_{CO}. This assumption is in agreement with literature data for CO adsorption on Pt group metals:

$$b_{fc} = b_{fc0}e^{\theta_{CO}\frac{\delta(\Delta G_{CO})}{RT}} \quad (36)$$

Variation of b_{fc} with θ_{CO} in Eq. (36), using a $\delta(\Delta G_{CO})$ value of 20 kcal mol^{-1}, reproduced the observed dependence of the ''critical current density'' (current at which the anode overpotential begins to steeply increase) on CO partial pressure for Pt/C anode catalysts. $\delta(\Delta G_{CO})$ is the difference in the adsorption-free energy assumed between $\theta_{CO} = 1$ and 0. Figure 33 shows the fits of such model calculations to experimental data for PEFCs operating with a range of CO levels in H_2 anode feed streams, duplicating the form of the polarization curves recorded, as well as the dependence of critical current on P_{CO} for CO levels in H_2 ranging between 25 and 250 ppm.

Basically, the ''critical current density'' feature is the result of a catalyst surface with much higher adsorption affinity to CO molecules than to H atoms (from H_2) and, at the same time, much higher electrooxidation rate of H_{ads} than of CO_{ads} in the relevant anode potential range. At such a catalyst surface, the overall rate of H_2 electrooxidation is determined by the limited rate of dissociative adsorption of hydrogen (25) taking place at a (very) limited number of anode catalyst sites remaining free of adsorbed CO. Since process (25) is a chemical, rather than electrochemical, step, that is, it is not associated with interfacial charge transfer, it cannot be accelerated beyond the ''critical current density'' by overpotentials typical for a CO-free hydrogen anode. Under such conditions, increase in anode potential can accelerate the rate of hydrogen electrooxidation only through electrochemical removal of adsorbed CO. This process (27) takes place at very low rates at anode potentials lower than 0.3 V versus hydrogen. To strongly accelerate process (27) further to secure a larger fraction of CO-free Pt surface sites and thereby a higher rate of (25) + (26), a large jump in anode overpotential, which results in the fall in cell voltage by 0.3–0.4 V seen in Figs 32 and 33, is required.

8.3.7.1.3 **Anode Catalyst Cleansing by Oxygen Bleeding** The cell behavior recorded in the presence of low levels of CO, as

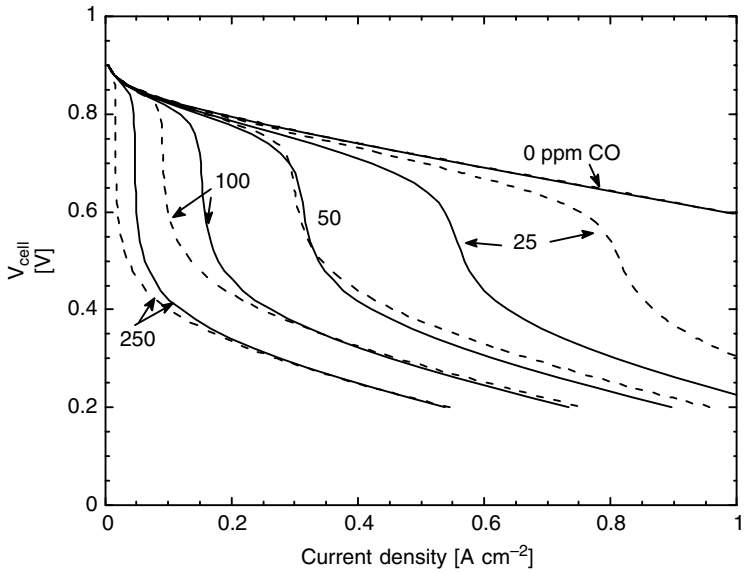


Fig. 33 Calculated cell voltage as function of current density based on the model described by Eq. (24–27). Dashed curves correspond to coverage-independent energy of adsorption of CO on Pt, whereas solid curves, which better fit experimental data, correspond to adsorption energy falling with coverage according a Temkin-type isotherm [42].

seen in Figs 32 and 33, clarifies the significant challenge of maintaining high performance in PEFCs operating with anode feed streams derived from processed carbonaceous fuels. When considering losses in the fuel-processing steps upstream the PEFC and, in addition, the significant cathode losses in the PEFC itself, it is readily understood that, to maintain an attractive energy-conversion efficiency, the PEFC cannot be allowed to operate at anode overpotentials significantly higher than 50 mV. Therefore, maximizing the ''critical current'' (Figs 32 and 33) at some given level of CO in the anode feed stream is essential. One possible approach demonstrated to the mitigation of CO poisoning at the PEFC anode catalyst is to bleed low levels of oxygen or air into the CO-contaminated anode feed stream [68, 69]. It was first demonstrated for Nafion-impregnated gas-diffusion anodes (0.4 mg Pt/cm^2) that bleeding of 0.4–2% oxygen into the CO-contaminated H_2 feed stream can completely correct for deleterious effects of 5–100 ppm CO. It is essential that the levels of oxygen employed in this in situ bleeding approach be well below the explosion threshold for O_2/H_2 mixtures (5% O_2 in hydrogen). Because of this limitation, the maximum level of CO that can be treated effectively by this in situ oxygen (or air)-bleeding approach is not more than several hundred parts per million. The effectiveness of the in situ bleeding of oxygen in removing deleterious effects of CO in PEFC anodes can be understood from the following sequence of processes at a PEFC Pt anode catalyst exposed to H_2 with residual levels of CO. As discussed earlier in this section, the affinity of CO to Pt surfaces at 80 °C is so strong that the

catalyst surface coverage by CO is very high in the presence of just 10–100 ppm CO. Schematically describing the case for, for example, 10 ppm CO in H_2, the Pt catalyst deactivation process can be schematically described as:

$$\begin{aligned}&(10\ \text{ppm})CO_{(@1\ \text{atm}H_2)} + Pt = Pt{-}CO_{ads}\\ &\quad(\text{also, } (10\ \text{ppm})CO_{(@1\ \text{atm}H_2)}\\ &\qquad + Pt{-}H_{ads} = Pt{-}CO_{ads} + 1/2H_2)\end{aligned} \tag{37}$$

"Cleansing" of the catalyst site by bleeding of dioxygen can then be described by the following equations:

$$\begin{aligned}&\text{(a) } Pt{-}CO_{ads} + (2\%)O_{2(@1\ \text{atm}H_2;10\ \text{ppm}CO)}\\ &\qquad + Pt = Pt{-}O_{ads} + Pt{-}CO_{ads}\\ &\text{(b) } Pt{-}CO_{ads} + Pt{-}O_{ads} = CO_2 + 2Pt\end{aligned} \tag{38}$$

with any oxygen beyond that consumed in process (38), reacting with hydrogen at CO-free Pt sites, according to

$$1/2O_2 + H_2 = H_2O \tag{39}$$

As Eq. (37) and (38) suggests, the strong preferential adsorption of CO at the Pt site is also the basis for its catalytic chemical oxidation by the O_2 molecules introduced into the anode feed stream. This *chemical* oxidation of CO by dioxygen at the Pt anode catalyst most probably takes place by dissociative chemisorption of oxygen followed by a bimolecular surface reaction between adsorbed CO and adsorbed O atoms (Eq. 38a,b) [70]. As described in the schematic reaction sequence, following oxidation of CO off the Pt surface, the remaining (i.e., most) introduced oxygen in the anode feed stream would recombine at Pt sites freed of CO, with hydrogen present in large excess in the anode compartment (Eq. 39). This recombination process results in loss of fuel (hydrogen) at a percentage that is twice the oxygen level in the anode.

Continuous bleeding of low levels of oxygen prevents anode poisoning from taking place altogether [68, 69] as the CO in the fuel feed stream continuously reacts at anode Pt sites with the oxygen added to the fuel feed stream according to

$$\begin{aligned}CO(g) + O_2(g)(+Pt) &= O/Pt + CO/Pt\\ &= CO_2 + Pt\end{aligned} \tag{40}$$

thus maintaining zero CO level in the gas phase within the anode compartment. This could be demonstrated experimentally by analyzing the anode exhaust stream when the anode feed stream contained 100 ppm CO and 2% O_2 was continuously added.

In the original demonstrations of such oxidative cleansing of the PEFC anode in the presence of up to 100 ppm CO [68], some excess of unutilized catalyst in the Nafion-impregnated gas-diffusion electrodes employed apparently provided sites for chemical oxidation of CO (Eq. 38). To further probe this understanding, a reactive "prefilter" for carbon monoxide, consisting of a Pt-catalyzed gas-diffusion electrode without any added ionomer, was introduced in the anode feed stream between the air bleed injection point and the anode inlet, thus adding more Pt sites in direct contact with the anode feed stream [71]. With this active prefilter in line, 1% air was sufficient for complete performance recovery at the level of 20 ppm CO in the H_2 feed stream, compared with 1% of neat O_2 necessary to take care of the same CO level without such an active prefilter. Finally, in light of the recently established understanding on the detrimental effects of $OH^{\bullet}$ radicals that tend to form by oxygen/hydrogen recombination at the PEFC anode (see Sect. 8.3.5.2), there is good reason to have the oxidative cleansing of the

hydrogen feed stream take place in a separate ''catalytic filter'' upstream the anode catalyst. With proper design of this active prefilter, very little oxygen would reach the anode catalyst.

8.3.7.1.4 **CO-tolerant PEFC Anode Catalysts** Although demonstrated successfully in PEFC stacks with CO-contaminated feed streams, air bleed into the anode is not considered a problem-free approach. Mixtures of hydrogen and oxygen are of some concern even in well-controlled systems. Consequently, the search for a PEFC anode catalyst of better CO tolerance has been pursued. At such an anode catalyst, the hope is that a sufficient rate of CO electrooxidation could be obtained closer to 0 V versus RHE, such that a sufficient number of ''free'' Pt sites could be maintained at steady state at low anode overpotential. This requirement corresponds to values of k_{ec} in Eq. (27) exceeding that for Pt catalyst by at least 2 orders of magnitude. Alternatively, or in addition, lower affinity of CO to an optimized anode catalyst surface, corresponding to higher b_{fc} in Eq. (24), would also bring about improved tolerance. PtRu catalysts have been most extensively researched, developed, and demonstrated in recent years as CO-tolerant anode catalysts in PEFCs [72–74]. Other alternative to PtRu anode catalyst exhibiting significant CO tolerance, for example, PtMo, have also been described [75]; however, such alternatives have not been fully demonstrated to have superior catalytic properties and long-term stability.

An approach described most recently for maximizing CO tolerance at relatively low anode catalyst loadings involves a combination of a CO-tolerant electrocatalyst, for example, a PtRu alloy and a separate chemical oxidation catalyst layer of zero ionomer content, embedded within the anode structure itself in front of the (ionomer rich) anode catalyst layer. The latter serves the same function as that of the ''catalytic prefilter'', originally described in Ref. 71. Such properly positioned, ''bifunctional'' anode catalyst structures allow to combine oxidative lowering of Pco in the anode by air bleed, with maximized tolerance of the electrocatalyst to any remaining low level of CO in the gas mixture reaching the electrocatalyst. PEFC performance level very close to that achieved on neat hydrogen was demonstrated with such a composite anode in operation with an anode feed stream of 40% hydrogen containing 100 ppm CO and is presented in Fig. 34 [42]. The figure also highlights the fact that CO_2, a major constituent in the anode feed stream in reformate/air PEFCs, is not playing a role other than that of fuel diluent. This is nontrivial because CO_2 could also react chemically with hydrogen (''reverse water shift'') at the anode catalyst to form CO in situ. Such reactivity of CO_2 in the PEFC anode was recorded with nonoptimized anode catalysts/structures.

8.3.7.1.5 **The PEFC Anode: Summary** The rate of hydrogen oxidation at an impurity-free, well-humidified Pt/ionomer interface with catalyst surface area of the order of 100 cm^2 Pt/cm^2 of membrane area is sufficiently high to limit anode voltage loss in PEFCs to several millivolts at 1 A cm^{-2}, as indeed observed with hydrogen feed streams of high purity. However, anode catalyst poisoning by CO is a major challenge that needs to be addressed in each PEFC-based power system according to the expected composition/residual contamination of the fuel feed stream. Levels of CO up to 100 ppm in the anode feed can be taken care of by in situ bleeding of air, using a properly

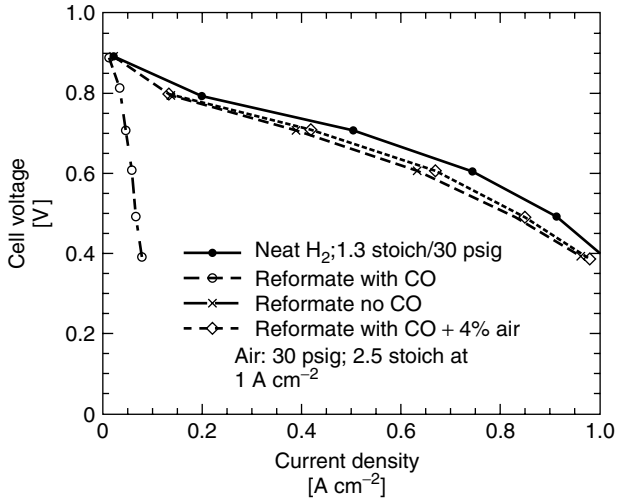


Fig. 34 Tolerance to CO, demonstrated experimentally at 70% fuel utilization with inlet fuel feed of 40% hydrogen containing 100 ppm CO. Bleeding air at 4% by volume into such anode feed stream is seen to restore the performance back very close to the CO-free, 40% hydrogen case. Precious metal loading in the cell anode was, in this case, 0.6 mg cm^{-2} [42].

configured anode. Such an anode would likely use a separate *chemical oxidation catalyst*, not in contact with the ionomer, placed apart from a CO-tolerant *electrocatalyst*, which is intimately mixed with the ionomer and placed in direct contact with the ionomeric membrane. While the "air bleed" approach has been shown to be effective in the operation of PEFC stacks with processed carbonaceous fuels alternatives to air bleed will undoubtedly be further pursued, at least because of any remaining safety concerns. One such straightforward alternative is relying on the best CO-tolerant anode electrocatalysts available, avoiding completely air bleed. This option, however, involves larger anode losses, particularly at CO levels of 100 ppm or above. Another future option considered for significantly enhancing the tolerance to CO, is based on the possible use of higher temperature membrane technology that would allow nonpressurized cell operation at temperatures in the range 120–150 °C. It is intrinsically difficult, however, to achieve high electrode performance in PEFCs operating at low water activities (see Sect. 8.3.7.2). Consequently, in membranes exhibiting high proton mobility above 120 °C, rendered, for example, by (neat) PA "dopant" [21] even at very low membrane water content – the cathode electrocatalytic activity is 2–3 orders of magnitude lower versus that achieved at 80 °C in PEFCs with hydrated poly(PFSA) membranes. To date, such trade-off between higher CO tolerance and lower cathode electrocatalytic activity remains the case for all membrane electrolytes tested at temperatures over 100 °C under low water activity conditions.

Finally, CO is not the only fuel-derived contaminant affecting anode performance

in the PEFC. In a test of other possible contaminants that could result from methanol reforming, for example, [76] strong and irreversible effects of formic acid at a PEFC were reported for a Pt-catalyzed (high loading) anode, whereas methanol, formaldehyde, and methyl formate were found to have much smaller and reversible effects. This is an example of a case where the use of a more active anode electrocatalyst, such as PtRu, could significantly facilitate oxidation of intermediate several significant impurities resulting from the coupling of natural gas (or gasoline) reformers of various types with PEFC stacks. The use of the latter fuels requires to probe and address, by either removal upstream or with modified catalysts, possible detrimental effects of low levels of sulfur, H_2S, COS, and NH_3. The effects of ppm levels of NH_3 in the fuel feed stream on PEFC performance have been recently reported in [77]. At the level of a few parts per million, the effects of ammonia on PEFC anode performance are major. They can be reversed by switching to perfectly ammonia-free hydrogen, provided the exposure to the ammonia is not longer than a few hours. After longer exposure, the deleterious effect becomes irreversible. Unlike CO, ammonia apparently does not deactivate the anode catalyst by adsorption at the ppm levels concerned and the deleterious effects seem to result from exchanging ionomer protons first in the anode catalyst layer(s) and subsequently also in the membrane by ammonium ions [77]. This lowers strongly the interfacial activity in the anode first and, with further exposure, the membrane proton conductivity. The clear conclusion from the information provided in [77] is that the level of ammonia has to be brought down to zero in the PEFC anode feed stream and that the target can be achieved, for example, by appropriate filtering.

8.3.7.2 The Air Cathode in the PEFC

8.3.7.2.1 PEFC Cathode Losses and the Uniqueness of the Pt Catalyst The high overpotential at the PEFC cathode is the single most important source of loss in the PEFC, as it is in other low-temperature fuel cells. The sluggish nature of the ORR process is immediately reflected by the measured open-circuit voltage of around 1.0 V, to be compared with the thermodynamically expected value of 1.23 V for a H_2/O_2 fuel cell (at 25 °C and unity activity of reactants and product). This discrepancy is caused by the very low rate of ORR (as well as that of water electrooxidation) at a Pt catalyst near 1.23 V, resulting in the PEFC open-circuit reading being a "mixed potential". The value of this mixed potential is determined by a cathodic ORR component of the zero net current at open circuit and an equal and opposite anodic component, which could be of several possible origins, for example, anodic oxidation of carbon from the catalyst support or from the backing material. A specific value of hydrogen/air PEFC open-circuit voltage is thus determined in the latter case by the requirement: $(J_{\text{oxygen reduction at Pt}})_{,\text{OCV}} = -(J_{\text{carbon oxidation}})_{\text{OCV}}$

A cathode potential of 0.7 V is required to reach an ORR current density of 1 A per cross-sectional square centimeter of a PEFC air cathode at 80 °C, corresponding to a cell voltage loss of 0.5 V. This is, by far, the largest loss in a hydrogen/air PEFC. Furthermore, such cathode performance requires Pt or Pt alloy catalyst of significant cost. The demonstrated 2006 PEFC cathode technology of 0.2 g Pt/kW_{peak}, translates, at the level of Pt

cost (900 $/oz) early in 2006, to cathode catalyst cost of nearly 10 $/kW$_{peak}$, that is, about 20% of the total target cost of PEFC stack technology in mass manufacturing for passenger vehicle applications. In fact, this is a state of PEFC cathode technology reached after very substantial developments between the mid-1980s and the mid-2000s, which led to lowering of the Pt loading in the PEFC by close to 2 orders of magnitude versus the loadings used in earliest PEFC (space) technology starting in the 1960s. Such R&D efforts to minimize the Pt loading were centered first at Los Alamos National Laboratory [7, 8, 10–12] and were followed subsequently by further advances made by fuel cell R&D teams at, for example, 3M and at GM [53, 78, 79]. Using carbon-supported Pt [7, 8, 10–12] and, most recently, Pt alloy catalysts on either carbon or on nanostructured ultrathin film support [53], the catalyst utilization of Pt catalyst in the PEFC air cathode was enhanced by almost 2 orders of magnitude versus state of the art in the mid-1980s. PEFC cathode Pt loadings were lowered from 4 to 8 mg Pt/cm^2 to under 0.1 mg Pt/cm^2 [53], while actually improving cathode performance (albeit, ultralow Pt loadings have not been proven as yet to provide sufficient longevity for some given long-term application). In each case, optimized catalyst layers of mixed electronic and ionic conductivity and of high reactant transport rates were critical in translating a given form and degree of Pt dispersion into a highly performing air cathode.

A probing question frequently brought up regarding PEFC technology is, "what is so unique about a platinum or Pt alloy oxygen reduction (ORR) catalysts that makes them irreplacable by less costly catalysts, possibly such catalysts that are not based on a precious metal?" The explanation for the uniqueness of Pt as an oxygen reduction catalyst in aqueous acid electrolytes is the optimized affinity of the Pt metal surface to oxygen. A Pt surface immersed in an aqueous acid electrolyte becomes significantly free of surface oxide species derived from water – see process (41) – at potentials just under 1.0 V versus the reversible hydrogen potential (RHE). At the same time, a Pt catalyst starts to exhibit significant reactivity at such high potentials versus the dioxygen molecule and, consequently, onset of the four-electron reduction of O_2 at the Pt metal – process (42) – also occurs just under 1.0 V.

$$PtOH_{ads} + H^+ + e = Pt + H_2O \quad (41)$$

$$Pt + O_2 + H^+ + e = Pt(O_2H)_{ads} \quad (42)$$

No other metal catalyst provides this match, in an aqueous acid electrolyte, between freeing metal surface sites at high V_{cath} (process 41) and allowing, at this high V_{cath}, onset of reductive adsorption of dioxygen (process 42), the first step in the overall four-electron reduction of dioxygen to water. Gold, a more noble metal than Pt, is free of surface oxide up to a higher potential of 1.2 V in acid solutions, but such low surface affinity to oxygen also results in much lower activity of gold versus dioxygen in acid electrolytes, making it a relatively poor ORR electrocatalyst in such electrolytes. On the other hand, metals of slightly higher oxygen affinity than Pt (e.g., Pd), start to exhibit oxygen reduction activity only at lower values of V_{cath} required to expose surface metal sites on Pd, resulting in larger ORR overpotentials. Metals of even higher affinity to oxygen (e.g., Ni) typically maintain a thin surface oxide layer in acid solutions through the complete relevant potential window, or are prone to electro- and, therefore, do not provide any

appreciable ORR activity in aqueous acid electrolytes.

That this ''optimized oxygen affinity'' is the key for the relatively high activity of Pt-based ORR catalysts has been further elucidated recently by the demonstrated enhancement in ORR activity achieved with some Pt alloy catalysts, for example, PtCr, PtNi, and PtCo [78–81]. Such an enhancement in ORR kinetics is understood to be the result of further fine-tuning of the surface affinity to oxygen, shifting the onset potential for process (41) slightly higher versus the case of unalloyed Pt. Thereby, the onset of the dioxygen reduction process (that requires surface metal sites) occurs at somewhat higher cathode potentials versus unalloyed Pt. Such slight lowering of the affinity to surface oxidation achieved by alloying Pt with active metals like Co, is understood to result from a core-and-shell structure of the Pt alloy particle [80], where surface-segregated Pt atoms that form the particle shell are bonded to core electropositive metal atoms, thereby lowering the tendency of Pt to surface oxidize (by the reverse of process 41).

Acceptability of some loading, hence cost of precious metal catalysts in a PEFC, is, obviously, highly application-dependent. Addressing this question directly for transportation applications, a GM fuel-cell team has targeted a level of <0.2 g Pt/kW [78, 79, 81]. It should be noted in this regard that recovery of precious metal catalysts from used PEFCs is a viable option and industrial level facilities for such recovery have, in fact, been implemented. The precious metal cost challenge is somewhat less daunting in stationary power applications and even less so in portable power. The catalyst loadings, in terms of gff/kW, required in the latter cases may actually be higher than in transportation because of the much longer life demanded in stationary power applications and the operation at lower cell temperatures in the case of portable power. However, acceptability of higher fuel-cell cost in the latter two applications derives from the higher cost per unit power of the incumbent technologies: several hundred dollars per kilowatt in the case of gas turbines for stationary power and several thousand dollars per kilowatt in the case of batteries for portable electronics.

8.3.7.2.2 **The Kinetics of the ORR at the Pt/Ionomer Interface** The reduction of O_2 occurs in the PEFC at the interface between dispersed Pt particles and a recast, hydrated ionomeric electrolyte. Important features of this electrolyte in the context of the ORR are high solubility of oxygen in the poly(PFSA) membrane ($\times 10$ the solubility of O_2 in water), immobilized anions, and a strong dependence of proton mobility on the local water content. The benefit of the higher solubility of the reactant is clear, as the ORR process is first order in oxygen partial pressure (described later in this section). Anion immobilization can minimize site blocking by anions that tend to adsorb on and deactivate Pt surface sites. On the other hand, the strong dependence of proton mobility on the local water content presents a challenge. It requires a high degree of ionomer hydration within the cathode catalyst layer to ensure a high rate of ORR and, at the same time, avoidance of excess liquid water buildup in the cathode to a degree causing flooding within the cathode-backing layer (GDL).

The overall voltage loss in the fuel-cell cathode is a complex combination of interfacial kinetics and mass and charge transport losses. Consequently, model interfacial systems are best suitable to resolve and directly study the ORR at the Pt/ionomer interface. Such model

systems require properties of (1) simple, well-defined geometry that enables accurate mass-transport corrections and (2) interfacial composition that mimics as closely as possible the interfacial environment within the PEFC cathode, that is, Pt in contact with hydrated ionomer and no added liquid electrolyte. Model systems of this type described in the literature have been based on two general schemes. In one, a smooth Pt electrode, or microelectrode, is coated by a film of recast ionomer to generate a Pt/recast ionomer interface. In the other, a Pt microelectrode is pressed onto an ionomeric membrane to generate a Pt/ionomeric membrane interface. To complete the electrochemical cell, counter and reference electrodes have to be in contact with the same ionic medium and this can be achieved by direct contact of these other electrodes with the ionomeric phase, or by contacting a reference electrode to the ionomeric phase through a ''salt bridge'' containing a dilute acid solution. The latter involves a negligible perturbation of the hydrous ionomer system.

In earlier attempts to study the ORR at the Pt/ionomer interface [82], a classical RDE configuration was used to control oxygen transport to a smooth Pt electrode coated by a recast film of Nafion. This configuration is required to complete the electrochemical cell by immersing the RDE in an aqueous electrolyte in which the other two electrodes are inserted, leaving some questions as to the possible effect of the liquid electrolyte on the measured interfacial rates. Results shown in Table 4, of mass-transport-corrected ORR currents at the ''bare'' Pt electrode and at the same RDE coated with a recast film of Nafion, indicate marginal enhancement of ORR rates at the lowest over potentials for the RDE coated with a recast film of Nafion. This marginal enhancement is noticeably smaller than expected from the documented $10\times$ higher solubility of oxygen in the ionomer.

Tab. 4 Measured ORR current densities for a ''bare'' and ionomer-coated RDE

V versus RHE	*Bare RDE in 0.5 M H_2SO_4 [mA cm^{-2}, 25 °C]*	*Filmed RDE in 0.5 M H_2SO_4 [mA cm^{-2}, 25 °C]*
0.95	0.056	0.11
0.90	0.27	0.34
0.85	0.74	0.71
0.80	1.8	1.1

Part of the explanation came from analysis of the actual concentration and diffusion coefficient of oxygen in the *recast film* of the ionomer. Resolution of the oxygen concentration and diffusivity revealed (Table 5) D and C values for recast, hydrated Nafion, which fall between aqueous solution and bulk ionomer levels.

A subsequent series of papers [83–86] described investigations of the ORR at a more a relevant interface between a Pt microelectrode and a humidified Nafion membrane, that is, without any electrolyte involved. The contact between the microelectrode and the membrane was achieved by mechanical pressure. In this case, the interface studied is Pt/*bulk* ionomer, rather than Pt/*recast* ionomer, and, therefore,

Tab. 5 Oxygen concentrations and oxygen diffusion coefficients in aqueous acid, recast Ionomer, and bulk ionomer phases [82]

	0.5 M H_2SO_4	*Recast film*	*Bulk Nafion*
$10^6 \times C$ (mol cm^{-3})	1.1	3.1	9–16
$10^6 \times D$ (cm^2 S^{-1})	18	4.0	0.1–0.4

it still does not mimic precisely the interfacial composition in PEFC cathode structures based on Pt/C//recast ionomer composites (see Table 5). Nevertheless, this arrangement enabled preparation of a Pt/ionomer interface and completion of the cell by other electrodes in direct contact with the membrane without use of any liquid electrolyte, employing a well-defined geometry: microelectrode of known diameter in direct contact with an ionomeric membrane. This geometry also enabled studies at elevated pressures and temperatures. Mass-transport-corrected results, derived from slow scan (2 mV sec^{-1}) measurements of ORR currents in such cell, are shown in Fig. 35. They are fitted with two linear Tafel domains, as has been traditionally done in the description of such ORR data. From such plots, the authors concluded by extrapolation an exchange current density, J_0, of $J_0 = 7.8 \times 10^{-7}$ A cm^{-2} associated with the domain of high Tafel slope (high polarization). The authors noted that the latter is ''the highest exchange current density reported for the ORR at room temperature at any Pt/electrolyte interface''. While the high apparent exchange current density may look interesting, it should be realized that direct comparison of *measured* ORR rates at the Pt/ionomer interface and at other Pt/dilute aqueous acid solutions does not reveal an exceptionally high rate at the Pt/ionomer interface [87]. The value of ORR exchange current densities obtained by extrapolation of experimental data to the equilibrium potential of the O_2/H_2O couple has been questioned in [88], suggesting that one should rely on comparative evaluations of ORR rates on current densities *measured* at some fixed V_{cath} of, or near, 0.9 V, rather than on current densities derived by extrapolation. This approach has been fully adopted in most recent evaluations of electrocatalytic activity in PEFC cathodes (see, for example, Refs 53, 81, and 89).

Work described in Ref. 86 includes a detailed investigation of the temperature dependence of ORR kinetic parameters at the Pt/bulk Nafion interface in the range from 30 to 90 °C. Plots of the log of mass-transport-corrected current density versus the potential of the Pt microelectrode showed an increase of a factor of about five in the rate of ORR at 0.90 V and about three at 0.85 V as the temperature was increased from 30 to 90 °C. The apparent activation energies for the two regions (low and high Tafel slope) were calculated from the dependence of the apparent exchange current density on T according to

$$E^{\#} = -2.303\, R[\mathrm{d}\log J_0/\mathrm{d}(1/T)] \quad (43)$$

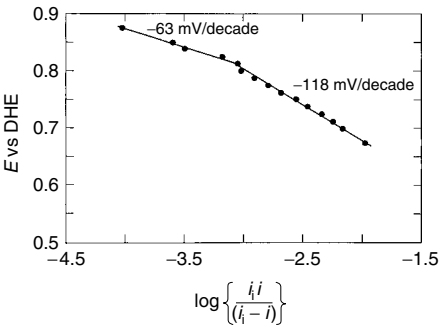


Fig. 35 Mass-transport-corrected $\log J_{cath}$ versus V_{cath} plot for oxygen reduction at the Pt/Nafion interface measured with the system described in Ref. 83.

yielding values of 17.5 kcal mol^{-1} for the low current density domain and 6.6 kcal mol^{-1} for the high current density domain. These are quite similar to activation energies of the ORR reported for Pt in contact with solutions of trifluoromethane sulfonic acid (TFMSA) or those in contact with TFMSA hydrate [86]. Additionally, the pressure dependence of the rate of the ORR was determined from the slope of measured log J_0 versus log Po_2 plots [86]. Unity slope was obtained for both neat oxygen and for oxygen in air, in both the low and high current density domains, as can be seen in Fig. 36. The first order of the ORR rate with respect to oxygen partial pressure applies to the Pt/ionomer interface, as it does to other Pt/aqueous acid systems, strongly suggesting that the rate-determining step is a one-electron charge-transfer step, usually written (although not as yet rigorously proven) as

$$O_{2,ads} + H^+ + e = (O_2H)_{ads} \qquad (44)$$

In a subsequent attempt to mimic more closely the composition and structure within the fuel-cell cathode ([87]), a *recast* Nafion™ *film* was applied to a Pt microelectrode in order to study the rate of the ORR as function of temperature and humidification conditions and the cell employed was as described in Fig. 37. This configuration comes closest to PEFC cathode conditions, in that the Pt surface is in contact with the recast ionomer, rather than bulk ionomer, and the atmosphere is humidified oxygen with no liquid electrolyte – only water vapor added. The voltammogram obtained for this filmed Pt microelectrode (in humidified inert atmosphere) is compared in Fig. 38 to the voltammogram recorded with the same electrode in aqueous sulfuric acid solution. For the filmed electrode, effects of site blocking are seen in the Pt–H domain and some anodic oxidation of surface impurities is apparent in the anodic half cycle above 1 V. This voltammogram did not change following several hours of cycling, suggesting some sustained loss of active surface area, which could be caused by site blocking by impurities

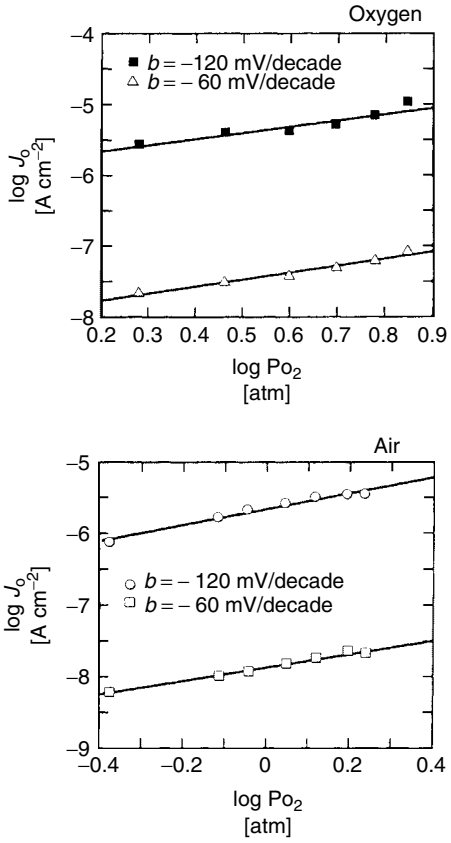


Fig. 36 Log J_0 versus log Po_2 plots for Pt cathode operating on oxygen and on air in contact with a Nafion membrane [83–86]. Extrapolations to the O_2/H_2O potential have been done from both the low current density/lower Tafel slope (60 mV/decade) branch and from the high current density/higher Tafel slope (120 mV/decade) branch [83–86].

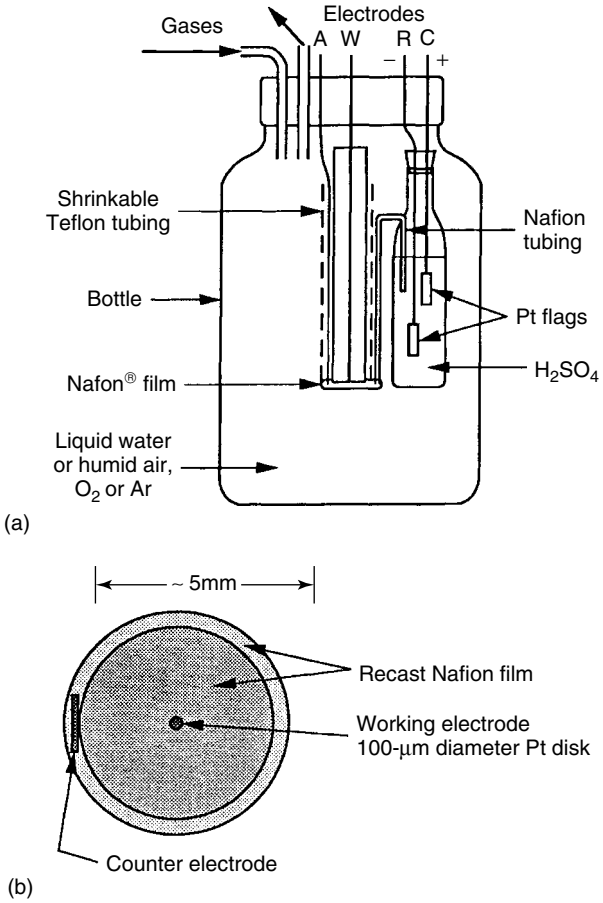


Fig. 37 Cell employed for measurements of ORR kinetics at the Pt microelectrode/recast ionomer interface (a) and a view of the microelectrode/counterelectrode geometry (b) [59].
W designates the cathode ("working") electrode, A, the anode, and R and C designate the dynamic hydrogen and the counterelectrode, respectively, for the DHE (dynamic hydrogen electrode) reference.

or by partial ionomer film delamination. Current–potential dependencies for the ORR process at such a Pt microelectrode coated with a recast Nafion film are shown in Fig. 39 for a range of temperatures. At each temperature, the gaseous atmosphere maintained around the microelectrode was saturated with water vapor. The results shown in Fig. 39 demonstrate the complex effect of temperature on the rate of the ORR at the Pt/recast ionomer in contact with water vapor. Against the expected beneficial effect of the temperature on the rate of the interfacial ORR process, a strong loss of ORR activity is actually seen above 40 °C for a filmed electrode in contact with saturated water *vapor*. This behavior is caused by loss of water

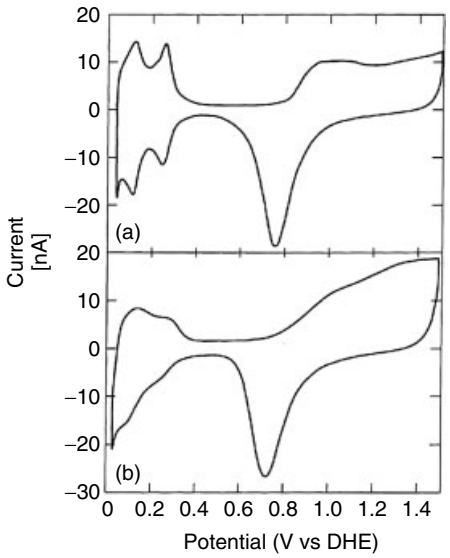


Fig. 38 Voltammograms obtained with the cell in Fig. 37(a) the microelectrode immersed in 0.5 M H_2SO_4 and (b) in contact with argon saturated with water vapor [59].

from the recast film, which takes place at elevated temperatures, although the water activity in the gas phase in contact with the film is maintained at unity (saturated vapor). A proof that drying of the ionomer caused this loss of ORR activity, was provided [87] by immersion of the filmed electrode temporarily in liquid water at 80 °C and showing that the ORR activity could be temporarily restored in this way. This was followed by renewed loss of ORR activity as the recast ionomer film gradually dried again while in contact with saturated water vapor at 80 °C.

The above observations explain an important aspect of ORR catalytic activity in PEFC cathodes. The sensitivity of the rate of ORR to interfacial water content and the tendency of the ionomer to dry up at elevated temperatures when in contact

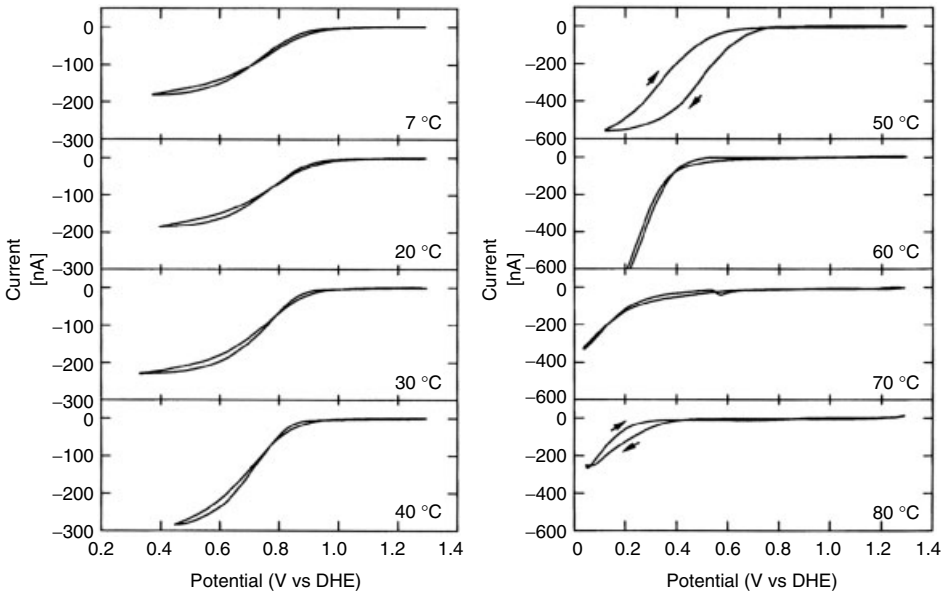


Fig. 39 Current–potential characteristics obtained with the cell in Fig. 37 at several temperatures in a water vapor–saturated oxygen atmosphere [59].

with saturated water vapor clearly suggest that, to ensure the highest level of cathode performance, the PEFC cathode catalyst should be kept in contact with some water in *liquid* state. This high local water activity requirement at the cathode catalyst/ionomer interface, was further elucidated in a subsequent PEFC impedance spectrum study [90] and ascribed to effective proton solvation being prerequisite for high rates of *interfacial proton transfer* (process 42). This specific need of good hydration of the cathode/ionomer interface, comes in addition to the requirement of high water activity in the bulk of the ionomer electrolyte to maintain high bulk membrane proton conductivity.

Mass-transport-corrected currents at $T = 30\,^{\circ}\mathrm{C}$, that is, the kinetic current densities evaluated from the measured $T = 30\,^{\circ}\mathrm{C}$ data in Fig. 36 are presented as a function of potential in Fig. 40, together with the corresponding value of $\mathrm{d}V_{\mathrm{cath}}/\mathrm{d}\log J$. The latter slope was computer-derived directly from the experimental plot of $\log J_{\mathrm{ORR}}$ versus V_{cath} [87]. Figure 37 shows a continuously varying Tafel slope through the cathode potential range of significance (0.9–0.5 V), with a current density of 0.1 mA cm^{-2} at 0.9 V recorded in the anodic scan (i.e., as the Pt electrode is not covered by excess chemisorbed oxygen). It was consequently suggested [87] that there is no good basis to ''force'' two discrete Tafel slopes on this data and that extrapolation of $\log J_{\mathrm{ORR}}$ versus V_{cath} data to the thermodynamic potential for the O_2/H_2O couple using some constant Tafel slope can yield exchange current densities for the ORR of questionable meaning.

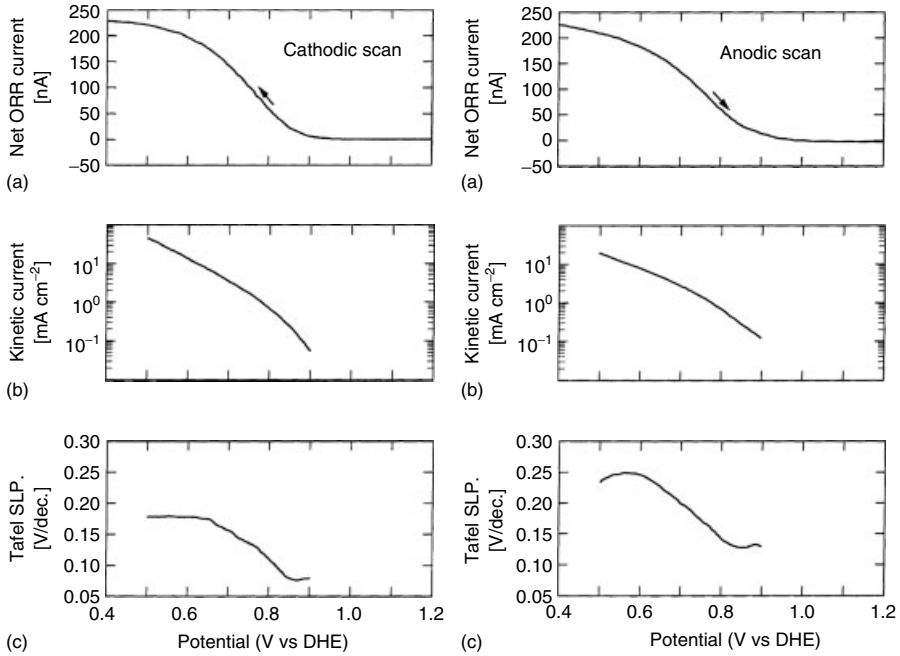


Fig. 40 (a) ORR current, (b) kinetic (mass-transport-corrected) ORR current and (c) the apparent slope $\mathrm{d}V/\mathrm{d}\log J$ as functions of cathode potential, obtained using the cell in Fig. 37 with water vapor–saturated oxygen atmosphere at 30 °C [59].

This traditional routine of extrapolation is questionable to V_{ORR} because the variation of the apparent interfacial ORR rate with applied potential at a Pt cathode is determined not just by the (usually considered) effect of cathode overpotential in enhancing electron transfer from the cathode to the dioxygen molecule (see Eq. 42) but also by the effect of cathode overpotential in removal of the surface oxygen species derived from water that attenuates the rate of the ORR. These two combined effects of cathode overpotential on the rate of the ORR result in an unusually small apparent $d\eta_{cath}/d\log J$ near $V_{cath} = 0.9$ V, increasing smoothly with cathode overpotential to reach eventually the ''ordinary'' Tafel slope expected for a single electron-transfer process like that shown in (42). The characteristics of ORR kinetic plots at Pt electrodes in acid electrolytes can be fully accounted for by the following simple model [88] based on the documented variation with a cathode potential of Pt metal site coverage by chemisorbed oxygen (from water). Assuming that (1) the rate of ORR at oxide-free Pt surface sites increases with overpotential according to a simple ''Tafel relationship'' at 120 mV/decade of current density and (2) the rate of ORR drops linearly with fractional surface coverage by chemisorbed oxygen derived from water (reverse of process 41), the mathematical expression for ORR current density takes the following form:

$$J_{ORR} = J_{V^c,\theta_{ox}=0}(1-\theta_{ox})\exp\left[\frac{(V^o - V_{cath})}{b_{int}}\right] \quad (45)$$

where θ_{ox} is the partial coverage by chemisorbed oxygen derived from water and b_{int} is the ''intrinsic Tafel slope'' of the ORR process at oxide-free Pt. Since θ_{ox} decreases with cathode overpotential, the change in ORR current with cathode overpotential is given by

$$d(\log J_{ORR})/d(V^o - V_{cath}) = 1/b_{int} + [1/(1-\theta_{ox})][d\theta_{ox}/dV_{cath}] \quad (46)$$

The first term on the right-hand side of Eq. (46) is the reciprocal of the ''intrinsic'' ORR Tafel slope, whereas the second reflects the effect of cathode overpotential on the rate of ORR brought about by further stripping of chemisorbed oxygen at more cathodic potentials. Figure 41 shows a calculated plot based on Eq. (46) employing $\theta_{ox}(V)$ measured by voltammetry for Pt in a relevant acid electrolyte solution and assuming that full coverage ($\theta_{ox} = 1$) by the chemisorbed (ORR blocking) oxygen species is attained at voltammetrically measured anodic deposition charge of 1 e per site. The calculated plot exhibits an initial branch around 0.95 V with an apparent very low slope of 40 mV/decade of current density, increasing gradually and reaching the intrinsic slope assumed of 120 mV/decade only as the blocking surface oxide has been completely stripped, around $V_{cath} = 0.75$ V. As the measured dependence of $\log J_{ORR}$ versus V_{cath} is seen to be determined by combined effects of b_{int} and $(d\theta_{ox}/dV_{cath})$, the only meaningful comparison between ORR rates obtained for different electrode/electrolyte interfaces should be based on mass-transport-corrected ORR currents *measured at a selected cathode potential*, rather than values derived by extrapolation. Such rates measured at 0.90 and 0.85 V versus RHE with the model system in Fig. 37 are shown in Table 6 for the Pt/recast ionomer and for some related, Pt/aqueous acid solution interfaces.

The ORR rates shown in Table 6 indicate an ORR activity at the Pt/ionomer interface near room temperature, which is quite

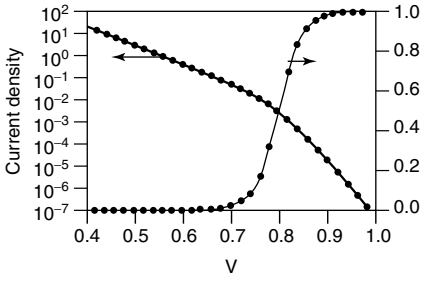


Fig. 41 Experimentally determined surface oxygen coverage, θ_{ox}, as a function of Pt electrode potential during a cathodic scan from 1.0 V and the associated $\log J_{ORR}$ versus V_{cath} dependence calculated (for ambient conditions) from the combined effects of enhancement of the rate of the ORR at Pt metal surface sites with increase in cathode overpotential and cathodic stripping of a blocking surface oxygen species between 1.0 and 0.6 V [88]. The $\log J_{ORR}$ versus V_{cath} plot shown, as calculated from Eq. 36, exhibits a gradual change of $dV/d\log J$ ("Tafel slope") from around 40 mV/decade at 1.0 V to 120 mV/decade at potentials lower than 0.65 V [88].

similar to that measured at interfaces of Pt with dilute aqueous acid solutions that do not contain strongly adsorbed anions. The activity at the Pt/recast ionomer interface is somewhat higher than that measured at the Pt/bulk ionomer interface, in spite of the lower concentration of oxygen in the former (Table 5) and the reason is probably the higher state of hydration of the recast film, apparently very beneficial for the enhancement of the rate of the ORR process at the Pt/ionomer interface (see Fig. 39). Such measured rates for the interfacial ORR process at the Pt/hydrated ionomer interface, as given in Table 6, can serve as basis for evaluation of catalyst utilization in the PEFC cathode.

Some ORR kinetic measurements reported most recently (2005) were performed directly on the dispersed catalyst of an operating PEFC cathode at an elevated temperature of 80 °C and a somewhat elevated air pressure of 1.5 atm [81]. The surface area of Pt available in such cathodes was evaluated by voltammetry, as shown in Fig. 42. The latter report adheres to ORR rate evaluation based on currents measured at 0.9 V versus RHE, and these results are given in Table 7. The iR-corrected, V_{cell} versus $\log J_{cell}$ plot in Fig. 42 shows a uniform slope of $dV_{cell}/d(\log J_{cell})$, however not that the plot is terminated near $V_{cath} = 0.8$ V (at 1 A cm^{-2}) and, consequently, the whole

Tab. 6 Measured ORR rates and apparent Tafel slopes at various Pt/acid electrolyte interfaces at 25 °C with gaseous atmosphere of O_2/saturated water vapor at $P = 1$ atm [87]

	mA cm^{-2} Pt at 0.90 V	***mA cm^{-2} Pt 0.85 V***	***b at 0.85 V mV/decade***
Pt/recast Nafion	0.1	0.4	80
Pt/bulk ionomer	–	0.2	63
Pt/0.5 M H_2SO_4	–	0.2	0.6
Pt/9.5 M TFMSA	0.2	0.4	118
Pt/6 M TFMSA	0.4	–	61

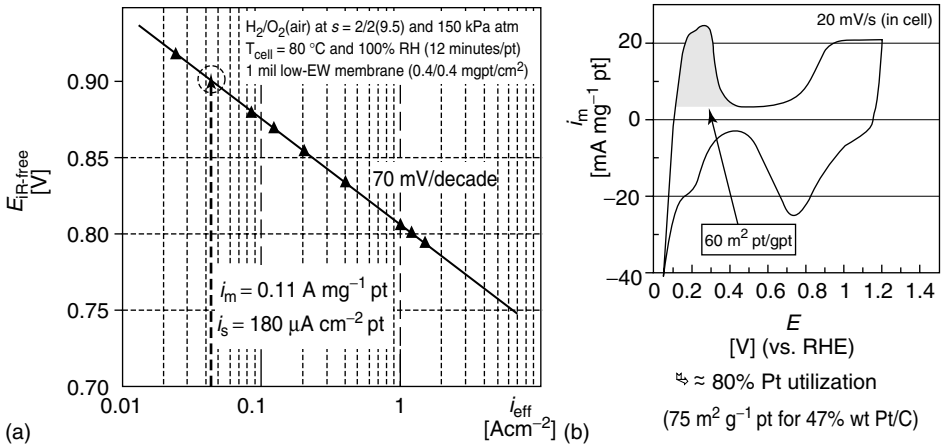


Fig. 42 O_2 reduction kinetics at a dispersed Pt Catalyst in a PEFC cathode operating at 80 °C; 100% RH (a). "Pt utilization", that is, Pt catalyst surface accessed by protons and electrons was determined for the cathode catalyst layer in question by cyclic voltammetry (b). ORR current per milligram of Pt at 0.90 V reflects state of the art for such cathode catalyst layers [81].

range of the cathode potential studied corresponds to potential dependence partial coverage by surface oxide, that is, to a lower Tafel slope (compare with Fig. 41). The activities of some commercial Pt/C catalysts measured in situ (Table 7) show ORR currents at 0.9 V cm^{-2} Pt at cell temperature of 80 °C and 1 atm O_2, about twice that measured at 25 °C for the model, flat Pt/recast Nafion system (see Table 6). According to the ORR rate enhancement expected with rise in temperature to 80 °C [86], the ORR rate measured at this elevated temperature should have been five times larger versus the rate at 25 °C. This discrepancy is explained by the documented loss of ORR activity per square centimeter of Pt when highly dispersed Pt particles (3 nm in diameter) are compared with flat surfaces of bulk Pt, or Pt particles of larger size [91]. Hence, the utilization of Pt in these cathodes is, in fact, as high as 90% [81]. On the basis of this recent, systematic testing, Table 7 also defines the

Tab. 7 Specific activity $i_{S(0.9\ V)}$ and mass activity $i_{m(0.9\ V)}$ of Pt/C catalysts in a PEFC cathode operating at 80 °C and ambient pressure of fuel and air: 2005 state of the art [81]

Catalyst	*20% Pt/C*[a]	*47% Pt/C*[b]	*46% Pt/C*[b]
An/cath [mg cm^{-2} Pt]	0.15/0.17	0.4/0.4	0.5/0.5
$A_{Pt,cath}$ [m^2 g^{-1} Pt]	71	75	86
$i_{S(0.9\ V)}$ [mA cm^{-2} Pt]	0.17	0.21	0.21
$i_{m(0.9\ V)}$ [A mg^{-1} Pt]	0.11	0.16	0.11

[a] Pt/C catalyst from source A.
[b] Pt/C catalyst from source B.

2005 state-of-the-art Pt/C activity in the ORR as 0.16 $A_{0.9V}$/mg Pt at 80 °C corresponding to 0.2 mA cm^{-2} Pt at 0.9 V at a catalyst dispersion level of 85 m^2 g^{-1} Pt (= 850 cm^2 mg^{-1} Pt). The latter is obtained with a Pt particle diameter of 3.5 nm. As Table 7 shows, the activity per milligram of Pt is not just maintained but even enhanced between the lowest and midlevel loading, apparently because of higher Pt dispersion in the Pt/C catalyst from source B.

Further increase in the loading to 0.5 mg Pt/cm^2, however, causes a drop in activity per milligram of Pt, as expected from transport limitations in the thicker catalyst layer (Sect. 8.3.7.2.3)

Further enhancement of the specific activity of PEFC air cathode catalysts has been achieved by moving on from carbon-supported Pt to carbon-supported Pt alloy catalysts [41, 78–81]. The gain in activity per unit mass of Pt in moving over from Pt to PtCo alloy cathode catalysts is demonstrated in Fig. 43 (explanation for this enhancement in activity has been given above). With the rise by factor 2.5 in ORR current per unit mass of Pt, achieved by the introduction of the PtCo catalyst, the coveted target of 0.2 g Pt/kW of PEFC power in transportation applications has been approached within about factor 2, albeit this is based on initial catalyst performance, that is, before long-term testing effects have been documented for such ultralow Pt loadings.

The possible complete replacement of Pt or Pt alloy catalysts employed in PEFC cathodes by alternatives, which do not require any precious metal, is an appropriate final topic for this section. Some non-precious metal ORR electrocatalysts, for example, carbon-supported macrocyclics of the type FeTMPP or CoTMPP [92], or even carbon-supported iron complexes derived from iron acetate and ammonia [93], have been examined as alternative cathode catalysts for PEFCs. However, their specific ORR activity in the best cases is significantly lower than that of Pt catalysts in the acidic PFSA medium [93]. Their long-term stability also seems to be significantly inferior to that of Pt electrocatalysts in the PFSA electrolyte environment [92]. As explained in Sect. 8.3.5.1, the key barrier to compensation of low specific catalytic activity of inexpensive catalysts by a much higher catalyst loading, is the limited mass and/or charge transport rate through composite catalyst layers thicker than 10 μm.

8.3.7.2.3 **The PEFC Air Cathode: Comprehensive Analysis of Voltage Loss** The overall voltage loss at the PEFC air cathode is determined by combination of interfacial and transport rate limitations, where

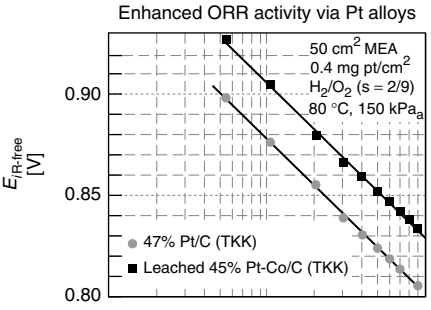


Fig. 43 Enhancement by factor 2.5 of the mass activity, (A mg^{-1} Pt)$_{0.9\ V}$, in the ORR process is achieved in the PEFC cathode by alloying Pt with Co [81].

the latter include charge and reactant transport in the composite catalyst layer and gas transport limitations in the wet-proofed backing layer (gas diffuser). To allow effective performance diagnostics, a model, based on physical and electrochemical parameters of the electrode, is required, providing a detailed description of all chemical and physical processes involved. Once such a model is validated against cathode polarization measured in the fuel cell for a range of cathode feed stream compositions, it can enable resolution of all factors contributing to an air cathode voltage loss at some given fuel-cell current density.

Such combined modeling/experimental diagnostics work for PEFC cathodes was described in [13]. The latter model addresses the cathode catalyst layer and gas-diffusion backing, as schematically presented in Fig. 44. The catalyst layer is considered in the model a composite film, typically 4–7 μm thick, of Pt/C catalyst intermixed with recast ionomer. The volume distribution of catalyst sites within this layer is considered uniform. The backing layer is a separate, wet-proofed carbon cloth (about 300 μm thick) of about 40% porosity. The effective porosity of the backing layer in the operating cell may become significantly smaller with increase in current of partial "flooding" by liquid water occurs. The cathode model in Ref. 13 considers losses caused by

- the rate of the ORR process at the Pt/ionomer interface,
- limited oxygen permeability and limited ionic conductivity within the catalyst layer,
- limited oxygen gas permeability through liquid-free pores in the wet-proofed backing layer, and
- drop in oxygen concentration along the air flow channel caused by oxygen reactant consumption.

The key model parameters that describe these types of loss in the PEFC are as

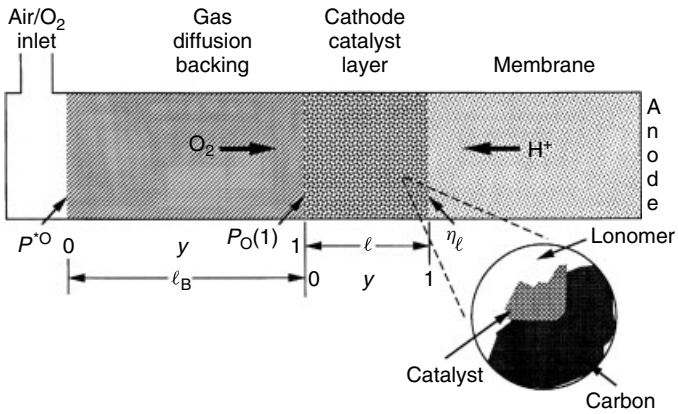


Fig. 44 Schematic of regions considered in comprehensive modeling of the PEFC air electrode. From left to right: gas flow channel, GDL, cathode catalyst layer. Oxygen in transported through the porous GDL by gas-phase diffusion through an inert mixture of nitrogen and water vapor. The catalyst layer is described in terms of effective transport characteristics of gas, protons, and electrons [13].

follows[13]:

- Interfacial ORR losses:

i^*, current density per square centimeter Pt at 0.9 V versus a RHE (A cm^{-2} atm^{-1}) at O_2 partial pressure $p_{ox} = 1$ atm;

a_r, the Pt surface area per unit geometric surface area (dimensionless); and

$\exp((0.9\,V)/b)$, or $\exp(F(V))$, the ORR rate dependence on potential relative to 0.9 V described by either a constant "Tafel slope" of 2.3 b (millivolts per decade of current density) or, alternatively, by an experimentally determined b(V) measured at a model, smooth Pt/ionomer system [88].

- Catalyst layer transport losses:

l_{cl}, the thickness of the catalyst layer (cm);

DC^*, the product of the effective diffusion coefficient of oxygen (cm^2 s^{-1}), and the equilibrium oxygen concentration (moles cm^{-3} atm^{-1}) in the catalyst layer at the interface with the backing (GDL),

$p_{ox} = 1$ atm;

σ, the effective protonic conductivity within the catalyst layer (S cm^{-1}).

- Backing layer transport losses:

l_b, the thickness of the backing layer (cm), and,

ε(J), the effective porosity of the backing (dimensionless).

The rest of the parameters that define the behavior of the cell for each set of operating conditions include the cathode inlet gas stream composition, flow rate and total pressure, the degree of humidification of the inlet gas stream, and the temperature of the cell.

The key equations describing the limitations in the catalyst layer in this model are as follows:

$$\frac{\partial \eta}{\partial y} = \frac{I \ell_{cl}}{\sigma} = I R_{cl} \tag{47}$$

$$\frac{\partial I}{\partial y} = p_{ox} i^* \frac{C}{C^*} a_r \exp\left[1/b_{cath}(\eta)(V_o - \eta)\right] \tag{48}$$

$$\frac{\partial C/C^*}{\partial y} = \frac{I - I_l}{p_{ox} I_D} \tag{49}$$

Equations 47–49 describe variations of parameters along the y coordinate of the catalyst layer ($y = z/l_{cl}$), where z is the catalyst layer thickness coordinate, $y = 0$ specifies the catalyst layer/gas interface, and $y = 1$ specifies the catalyst layer/ionomeric membrane interface (see Fig. 44), in which R_{cl} ($= l_{cl}/\sigma$) is the protonic resistance through a unit cross-sectional area of the catalyst layer and I_D ($= nFDC^*/l_{cl}$) is a characteristic diffusion-controlled current density in the catalyst layer. The thickness of the catalyst layer disappears from the equations by introducing R_{cl}, a_r, and I_D. The experimental variables considered include the overpotential η, the current density I, and the oxygen concentration C^*, when $p_{ox} = 1$ atm at the catalyst layer/gas interface. The O_2 partial pressure, p_{ox}, at the catalyst layer/backing layer interface is determined, in turn, by the cathode inlet gas stream composition and stoichiometric flow rate and by the backing layer (GDL) transport characteristics.

Equations (47–49) describe the nature of the losses incurred by transport within the catalyst layer. The catalyst surface area (catalyst sites) active in the ORR process of "conversion of oxygen flux to protonic current flux" (Eq. 49) is distributed linearly along the thickness of the catalyst layer. Consequently, gas supply to a catalyst site removed from the backing layer/catalyst layer interface is limited by the effective permeability of oxygen in the catalyst layer,

whereas proton reach to a catalyst site removed from the ionomeric membrane is limited by the effective protonic conductivity of the catalyst layer. As result of the former, the catalyst sites experience a distribution of oxygen concentrations (Eq. 48, 49), whereas, as result of the latter, different catalyst sites experience different effective overpotentials (Eq. 47). Calculated distributions of reactant and of electrode potential within the catalyst layer are described in Fig. 20 for the case of diluted hydrogen feed streams and similar types of distributions for the case of an air cathode were reported in Ref. 13.

Figure 45 shows a recent three-dimensional computer calculation [38] of the contours of oxygen concentration within the volume of a 10-μm thick composite catalysts layer, corresponding to the case of a PEFC operating at currents that are a significant fraction of I_D (Eq. 49). This recent modeling of the reactant profile in the catalyst layer was done using numerical reconstruction by a statistical method of the layer's actual microstructure, based on a series of 2-D micrographs [94]. Under conditions depicted by Fig. 45, the oxygen concentration is seen to fall in the part of the catalyst layer most removed from

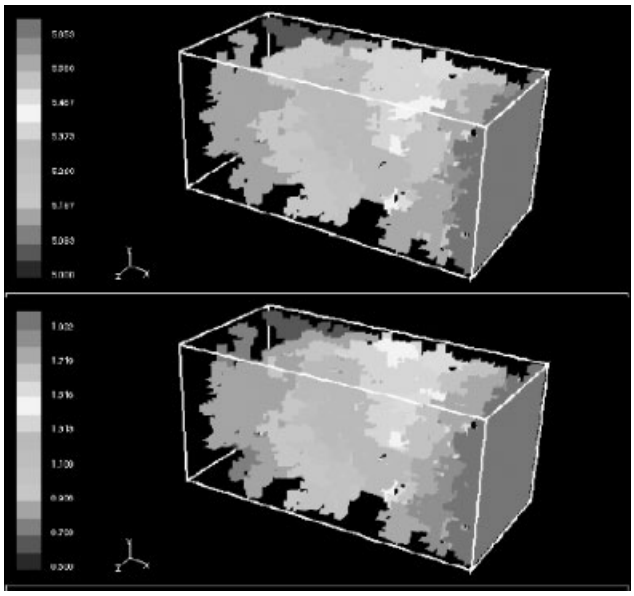


Fig. 45 Calculated 3-D contours of the oxygen concentration in a 10-μm thick cathode catalyst layer of a PEFC at 0.5 A cm^{-2} (a) and 1.5 A cm^{-2} (b) [94]. The calculation was done using numerical reconstruction by a statistical method of the layer's microstructure. The drop in oxygen concentration across the thickness dimension of the catalyst layer is 20% in the top case (J_{cell} = 0.5 A cm^{-2}) and 80% in the bottom case (J_{cell} = 1.5 A cm^{-2}). The concentration of oxygen at the backing/catalyst layer interface (designated by red) corresponds in the top and bottom case respectively to 70% and 20% of the concentration of oxygen in air; this drop is caused by the oxygen concentration gradient across the cathode GDL at some J_{cell}.

the backing layer, that is, most removed from the surface of oxygen entry. The relative drop in oxygen concentration across the thickness dimension of the catalyst layer (i.e., the ratio of concentrations designated by grey shading in Fig. 45) is 1.2 in the top case ($J_{cell} = 0.5$ A cm^{-2}) and 4 in the bottom case ($J_{cell} = 1.5$ A cm^{-2}). As ORR interfacial kinetics is first order with oxygen concentration, this drop in concentration in a significant part of the catalyst layer would cause inefficient catalyst utilization at higher current densities, resulting in further cathode voltage loss. Note also that the concentration of oxygen at the backing/catalyst layer interface (designated by red) corresponds in the top and bottom case respectively to 70% and 20% of the concentration of oxygen in air; this drop in oxygen concentration at the surface of entry into the catalyst layer is caused by the oxygen concentration gradient across the cathode GDL that scales with $J_{cell}/J_{lim,cath}$. In the case described in Fig. 45, $J_{lim,cath}$ is 2.1 A cm^{-2}, as set by the oxygen permeability of the backing layer (see Fig. 44). From Eq. (49), the way further improve high catalyst utilization at such high cell currents is to maximize I_D, which, in turn, requires a catalyst layer as thin as possible to contain all the catalyst surface area required to generate cell current. Comparison of results of such 3-D computations [94] with those of the 1-D macrohomogeneous model described in Ref. 13, allow to evaluate the effective medium key parameter (''Bruggeman factor'') for the macrohomogeneous model. For example, for proton mobility in a catalyst layer containing ionomer uniformly distributed at 11% by volume, a 3-D calculation based on detailed layer microstructure [94] indicated high tortuosity of close to 20 for the proton path, requiring use of a Bruggeman factor of 3.5 in the macrohomogenous model. It can be readily understood (albeit not always readily implemented) that the key to cathode catalyst layer optimization is simultaneous provision of high transport rates of oxygen, protons, and electrons, which, in turn, requires optimized catalyst layer volume fractions of ionomer, electronic conductor, and void volume. And, the thinner the catalyst layer containing the needed active catalyst surface, the lower is the transport loss in the layer.

The backing layer is treated in Ref. 13 with the Stefan–Maxwell equation to yield the partial pressure of oxygen at the backing/catalyst layer interface from the total pressure P_{tot} in the gas flow channel, the backing characteristic current density, I_B, and the mole fractions in the gas flow channel of water vapor and of oxygen, x_{ws} and x_{ON}, respectively:

$$p_{OX} = P_{tot} \cdot (1 - x_{ws}) \cdot \{1 - (1 - x_{ON}) \exp[I/I_B]\} \quad (50)$$

The ''backing characteristic current density'' is given by

$$I_B = nF D_{SON} \varepsilon^{1.5} (T/T_S)^{0.823} / V_m l_B \quad (51)$$

and is a function of the standard molar volume of an ideal gas, V_m, the oxygen/nitrogen binary diffusion coefficient, D_{SON}, at standard conditions, the backing thickness, l_B, and the backing porosity, ε. The effective porosity, ε, of the GDL and the way it is affected by the accumulation of liquid water enables evaluation of the limiting current of the cell (Ref. 90 considers, in addition, the effect of water vapor flax in the cathode backing and adds, based on impedance spectra analysis, effects of tortuosity of the oxygen path in the backing layer).

As described in Ref. 13, to obtain confidence in a model based on so many cathode parameters, simultaneous fit of a family of iR-corrected polarization curves, recorded for the same cell with a wide range of cathode feed streams, composition and total pressure was performed. Such a fit is shown in Fig. 46 and the captions provide the parameters, which yielded the best fit to the model. Two aspects to highlighted from the simultaneous fit of the four polarization curves shown in Fig. 46 were (1) oxygen transport rate in the gas phase within the porous backing layer determines the cell limiting current in operation on air and, (2) protonic conductivity limitations in the catalyst layer affect the performance in the midcurrent range. Curves (c) and (d), obtained for diagnostics using diluted oxygen feed streams, demonstrate that the PEFC-limiting current depends on oxygen mole fraction, rather than on oxygen partial pressure. This is immediate indication that transport in the air cathode is limited at high current densities in the *gas-phase* component of the porous GDL. On the other hand, comparison of curves (a) and (b) demonstrates the low enhancement in current observed at low overpotentials when changing cathode feed from 5-atm air to 5-atm O_2. Instead of near fourfold increase in current expected under control by interfacial kinetics (1st order in P_{O_2}) the observed increase at low cathode overpotentials (high cell voltage) is only about twofold. This ratio could be rationalized only by the effect of limited protonic conductivity within the catalyst layer, the relative importance of each effects varies

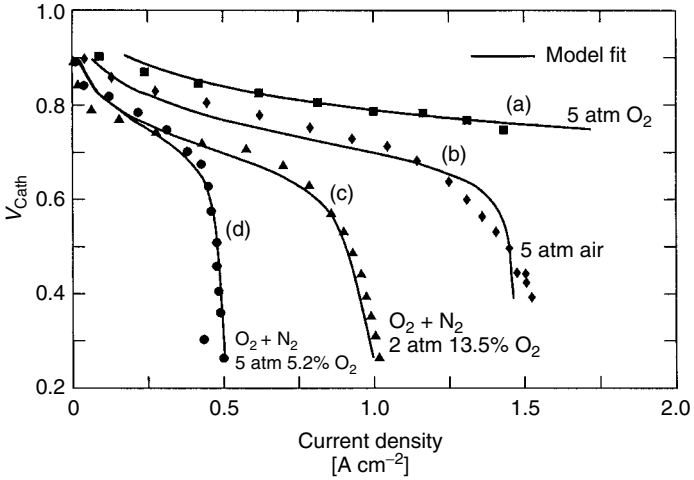


Fig. 46 Simultaneous fit to four polarization curves for a PEFC with 300-mm-thick cathode GDL and 7.5-μm-thick cathode catalyst layer applied to the membrane surface [12]. Different cathode feed stream compositions were used and a simultaneous fit demanded for all characteristics using the same physical and transport parameters for the GDL and the catalyst layer [13]. Fitted cathode parameters were: $i^*_{cath} = 90$ mA/cm^2, $b = 85$ mV/decade, $G = 1.3 \times 10^{-3}$ S/cm, $DC^* = 1.9 \times 10^{-8}$ mol/cm atm, and GDL porosity ε, varying between 0.25 at zero current and 0.19 at 15 A/cm^2.

with cell-operating conditions. For example, protonic conductivity limitations in the catalyst layer were found particularly important under cathode operation with low air pressure (see top part of Fig. 19).

The model in Ref. 13 suggested the following general trends:

- Gas-phase transport limitations in the air cathode GDL determine the cell limiting current and can affect the slope of the polarization curve in the medium current density domain.
- Transport limitations originating from the catalyst layer can be significant in the medium current density domain, particularly so at low air pressures.
- For neat oxygen feed, where backing effects are practically eliminated, PEFC losses from ionic resistance and permeability of the catalyst layer prevail.

Finally, behavior indicative of ''partial flooding'' of the fuel cell was modeled in Ref. 13 by assigning diminished porosity to the backing with cell current and inclusion of this feature was a prerequisite for complete modeling at high PEFC current densities.

In a more recent work [38], details of the transport of liquid water were considered in a system comprising gas channels adjacent a porous gas-diffuser (backing) layer. A key issue at the gas channel/backing layer interface (left most interface in Fig. 41) is the mode and rate of transfer of liquid water from the GDL to the airstream at some given rate of airflow in the channel. To help with the understanding of that aspect in an operating PEFC, direct observation of liquid water in the cathode backing and in the gas channel as function of position along the active area has been pursued. A transparent cell described [95] enabled direct observation of liquid water droplets coalescing in and emerging from the cathode backing into the airflow channel, as shown in Fig. 47. The effectiveness of liquid droplet transfer from the backing to the channel was reported to improve by rendering the channel surface hydrophilic properties, contrasting the hydrophobic nature of the backing material.

Observation of water distribution within an ordinary (unmodified) PEFC was subsequently enabled by application of a neutron beam probe, which allows to examine the water profile in an operating, ordinary cell without interference from the graphite hardware. Such image, recorded in an operating PEFC, is given in Fig. 48. The cathode inlet is at the upper right of the image and the exit is at the lower left; the anode inlet is at bottom left and the exit is at the upper right. Preferential liquid accumulation observed in the 90° bends in the flow field is typical, and

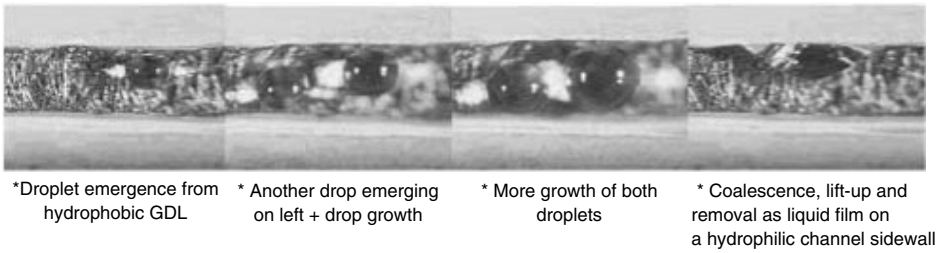


Fig. 47 Microscopic imaging of liquid droplet emergence and growth within a hydrophobic cathode-backing layer, followed by its removal into a hydrophilic air channel [38].

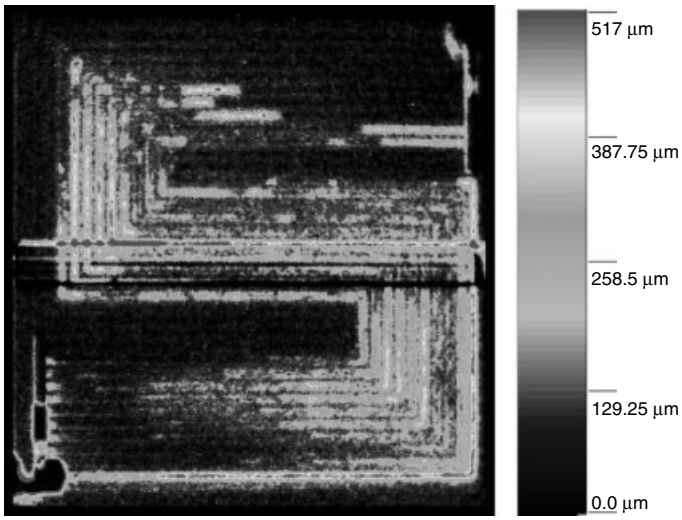


Fig. 48 Neutron image of an operating PEFC showing nonhomogeneous distribution of liquid water content [96].

is understood to be associated with lowered gas velocity at bends and corners. This image is of a 50-cm^2 PEFC under steady-state current of 0.4 A cm^{-2} (100% RH feed streams). The total liquid water mass accumulated in the cell-backing layers and channels could be quantitatively determined from this image, as 12 mg per cm^{-2} of the cell active area. The channel to interchannel ("land") water distribution could also be measured from this profile as being around 50/50, that is, the lateral distribution of the accumulated liquid through the backing layer is sufficiently effective such that liquid accumulation is not confined to the parts of the backing right under the channel. Since a typical volume of gas channels + backing layers in a PEFC is around 0.2 cm^3 per cm^{-2} of active area, the 0.01 cm^3 of water accumulated per square centimeter of active area does not amount at 0.4 A cm^{-2} to more than a few percentage points of the total open volume in the cell. However, significant localization of the same amount of liquid water at the backing GDL/catalyst layer interface could amount to a significant reduction in the rate of oxygen supply (the neutron beam probes the integral amount of liquid along the cell thickness dimension at each point along the surface and it does not provide resolution along the thickness dimension). Note that liquid water storage in the cell under steady state is a strong function of GDL material and geometry, of channel and "land" geometry and dimensions, and of operating conditions including cell current and RH of the feed streams. Therefore, the degree of liquid water accumulation shown in Fig. 48 does not represent a general case.

8.3.7.2.4 **PEFC Cathode Modeling and Diagnostics Based on Impedance Spectra**
Measurement and analysis of ac-impedance spectra of PEFCs, describing the generic form of spectra measured for such fuel cells and their value for cell diagnostics, was first provided in Ref. 90. The measurements were carried out using

instrumentation enabling the determination of the impedance spectrum in the presence of large dc currents. The impedance spectrum of the air cathode was shown to contain two features: a higher frequency loop determined by interfacial charge-transfer resistance and catalyst layer properties, and a lower frequency loop determined by gas-phase transport limitations in the GDL. The lower frequency loop is completely absent from the spectrum of cathodes operating on neat oxygen. A typical family of impedance spectra measured for PEFCs with air cathodes is shown in Fig. 49 and these spectra were analyzed by adapting the PEFC cathode model described in Ref. 13 for effects of ac perturbation. Comparison of model predictions to observed data was made by simultaneous least squares fitting of a set of spectra measured at several cathode potentials (see Fig. 49). The results revealed, for the specific PEFC cathode used, significant ionic transport and oxygen transport limitations in the cathode catalyst layer and significant gas transport limitations in the hydrophobic cathode backing. Both findings were reported before from steady-state polarization data and modeling [13]. Two features of the fuel-cell cathode were, however, uniquely highlighted by the impedance spectra: the multifaceted effect of cathode dehydration and the tortuosity of the oxygen path through the backing layer. Three different types of losses caused by insufficient cell hydration could be clearly resolved in the impedance spectra measured [90], related with slower cathode interfacial kinetics, loss of proton conductivity in the catalyst layer, and drop in membrane conductivity. Such a resolution would be practically impossible, to achieve in steady-state measurements.

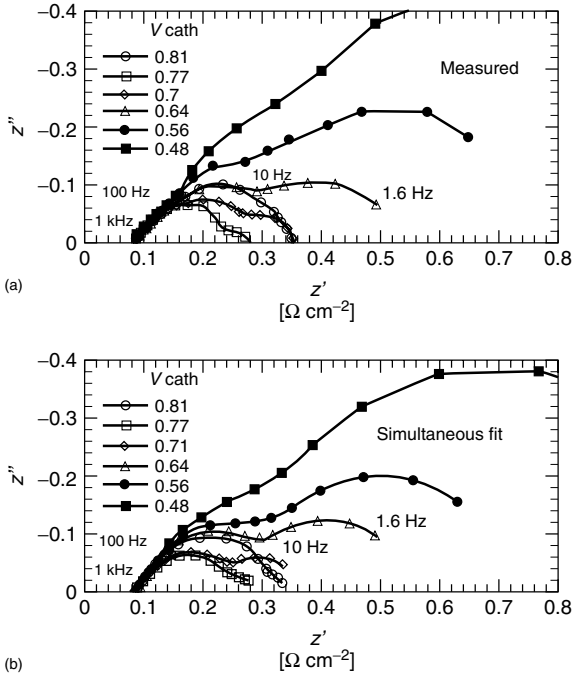


Fig. 49 Measured (a) and fitted (b) impedance spectra for a PEFC air cathode in a range of cathode potentials between 0.81 and 0.48 V versus hydrogen [90].

8.3.7.2.5 **The PEFC Cathode: Summary** To a large degree, the success of PEFC technology to date has been based on successful development of air cathodes that can exhibit, under minimal overpressures, rates (at 80 °C) of the order of 1 A cm^{-2} at $V_{cath} > 0.7$ V, using Pt loadings as low as 0.4 g Pt/kW [81]. Considering the very low rate of the ORR process per square centimeter of Pt in acid aqueous systems, a critical prerequisite for achieving this impressive state of PEFC air cathode technology has been maximizing and stabilizing of catalyst dispersion and optimizing the thickness and composition of catalyst layers. Another important factor has been the optimization of the cathode backing (GDL), which carries the key for operation under sufficient MEA humidification levels while preventing cathode flooding by excess liquid water accumulation next to the catalyst layer. Recently, a successful demonstration of Pt alloy catalysts, particularly PtCo, that exhibit activity per unit mass of Pt, which is two to three times higher than that of unalloyed Pt, is encouraging in suggesting further remaining options for fine-tuning PEFC cathode catalyst characteristics. It is also true that, with all the impressive recent achievements in this area of science and technology, the air cathode continues to be the single largest source of PEFC losses, caused by the need of a cathode overpotential or cell voltage loss of 0.4–0.5 V for accelerating the ORR process to current densities typically demanded by the application. To date, the possibility of enhancing the rate of the process by raising cell temperature has run against the high sensitivity of the ORR kinetics to interfacial water content, resulting in need of pressurized PEFC systems to just maintain ORR activity over 100 °C at the level observed in PEFCs with hydrated membrane electrolytes at 60 °C. Is a totally unexpected ORR catalyst for acid aqueous systems still hidden out there and waiting to be discovered? The likelihood is not high. However, further advancement by use of catalyst alloy nano-structures and compositions, along the lines of the recent success with Pt alloys of core and shell structures [78], is more likely.

8.3.8 The Direct Methanol Fuel Cell

The DMFC operates, as its name suggests, by direct, complete electrooxidation of methanol to CO_2 at the cell anode. The methanol anode is coupled in the DMFC with an air cathode, completing a cell schematically shown in Fig. 50. In the majority of recent development efforts, the DMFC has been based on a proton-conducting polymeric membrane. The uniqueness of this type of fuel cell is the direct anodic oxidation of a carbonaceous fuel. Such a direct electrochemical conversion process of liquid fuel and air to electric power at low temperatures can provide a basis for a very simple fuel-cell system.

Until very recently, only two fuels have been documented to undergo direct and complete electrooxidation at the anode of a fuel cell at ambient or slightly elevated temperatures: hydrogen (including some hydrides, for example, $NaBH_4$, that can be directly electrooxidized from an aqueous solution) and methanol. Compared to the case of hydrogen fuel, there is a significantly higher demand on the anode electrocatalytic activity when a carbonaceous fuel is to be directly and completely electrooxidized at a temperature under 100 °C. Unique properties of the methanol molecule that enable sufficient anode activity under such conditions are (1) the

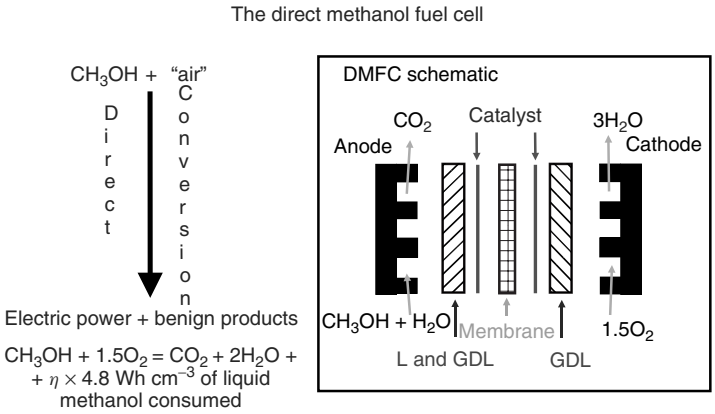


Fig. 50 Schematic presentation of the basic components and processes in a DMFC.

single carbon atom in the molecule (i.e., the absence of a carbon–carbon bond) and (2) the state of partial oxidation of the fuel molecule, that is, the presence of an oxygen atom in the molecule setting an atomic ratio, O:C = 1. The presence of a single carbon atom in the fuel molecule removes the need of carbon–carbon bond activation on route to complete electrooxidation and the significance of this aspect can be readily judged from the conditions reported for electrooxidation of ethanol to CO_2 (and some other side products) using a similar, Nafion-based PEFC and similar electrolytes: the temperature required is 160 °C, achieved in such a cell by operation under a pressure of 5 atm [97]. The carbon–carbon bond activation, required to oxidize the ethanol molecule all the way to CO_2, seems to set an activation barrier much higher than that required for complete electrooxidation of the single carbon molecule, methanol. As to the percentage oxygen in the molecule as key for ease of direct electrooxidation, a recent report sheds interesting light on this aspect. Peled and coworkers have shown [98] that some fuel molecules with a carbon–carbon bond can, in fact, be fully electrooxidized to CO_2 at low temperature in an acid electrolyte. The specific molecules for which they demonstrated such capability were oxalic acid, oxalates, and ethylene glycol. The common feature for all three is the high, 1:1 oxygen-to-carbon atomic ratio in the molecule (this ratio is 1:2 in ethanol). This high degree of ''oxygenation'' is apparently the key for enabling electrooxidation of a two-carbon atom molecule all the way to CO_2 at low temperatures. Cells based on any of the above two-carbon fuel molecules could consequently be further developed for some applications. However, at a given temperature of 60 °C or below, their reported power output is inferior to that of the DMFC [98].

In spite of the significant ease of anodic electrooxidation when compared with any other carbonaceous fuel, methanol electrooxidizes at significantly lower rates than hydrogen. With state-of-the-art polymer electrolyte DMFC technology at an anode overpotential near 0.3 V (versus hydrogen reference), the rate is similar to that exhibited by a hydrogen anode at an overpotential of several millivolts. Consequently,

similar cell current densities are obtained in DMFCs at cell voltages about 0.3 V lower than in hydrogen/air cells. As result, the central challenges in DMFC science and technology are (1) to achieve sufficient cell/stack power density and efficiency in spite of "anode inferiority" and (2) to pursue very simple DMFC system designs. A DMFC power system of >20% conversion efficiency achieved with DMFC stack power density of 100 W L^{-1}, or higher could allow to take full advantage of the attractive properties of methanol fuel – a liquid fuel under ambient conditions of specific energy significantly higher than that of relevant forms of hydrogen storage.

8.3.8.1 The DMFC Anode

The process of direct anodic oxidation of methanol at low temperatures is the basic enabler of the coveted low-temperature fuel-cell system based on direct feed by liquid fuel. This DMFC anodic process was only slowly recognized to have a real potential for continuous generation of current. A very common prevalent perception has been that the process is "self-poisoned" because an adsorbed CO intermediate originating from partial oxidation of methanol is gradually blocking and deactivating the catalyst surface. Much of that perception is due to the use of (unalloyed) Pt catalysts in researching the methanol electrooxidation process. These Pt catalysts have only one of the two requisite surface catalytic features for effective electrooxidation of methanol – they are very effective in facilitating the activation of C–H bonds in the methanol molecule and the subsequent fast electrooxidation of the resulting adsorbed H atoms. Pt is not the proper choice, however, for the other function required from a catalyst for complete methanol electrooxidation: electrooxidation of the CO intermediate, which indeed remains adsorbed on a Pt site after the C–H bonds in the methanol molecule are effectively cleaved and hydrogen atoms readily oxidized. To remove the CO intermediate by electrooxidation it at relatively low anodic overpotentials to CO_2, a catalyst of much higher surface affinity to oxygen than Pt is required. Pt catalysts start forming active surface oxygen from water in acidic systems at potentials over 0.6 V versus RHE – much too high an overpotential for a fuel-cell anode. The same process takes place at more oxophylic metals, like Ru, at much lower anodic overpotentials, around 0.3 V versus RHE [99]. From the above description of the reactivities of Pt metal and of Ru metal surfaces, it becomes clear why PtRu alloy catalysts have ended up the most active DMFC anode catalysts to date, being capable of addressing both required steps in the complete electrooxidation of methanol, that is, C–H bond activation and CO electrooxidation, at relatively low overpotentials [99].

The steps comprising the complete DMFC anodic process can be described according to the following scheme:

$$CH_3OH + 5Pt = 4Pt{-}H_{ads} + Pt{-}CO_{ads} \quad (52)$$

$$4(Pt{-}H_{ads}) = 4Pt + 4H^+ + 4e \quad (53)$$

$$Ru + H_2O = Ru{-}OH_{ads} + H^+ + e \quad (54)$$

$$Ru{-}OH_{ads} + Pt{-}CO_{ads} = CO_2 + H^+ + e + Pt + Ru \quad (55)$$

The overall anode process

$$CH_3OH + H_2O = CO_2 + 6H^+ + 6e \quad (56)$$

thus depends on Pt surface sites on the alloy surface to activate the −C−H bonds (and −O−H bond) in the molecule

and oxidize the hydrogen atoms derived from them, leaving an adsorbed CO intermediate on an adjacent Pt site, and further depends on Ru sites on the alloy surface to activate water molecules to form an active surface oxygen species (''OH intermediate''), which reacts with the adsorbed CO intermediate to complete the anodic oxidation all the way to the CO_2 product. Significantly, depending on whether the earlier step (52) or the later steps (54) + (55) become rate determining, the overall anode process could be first order or zero order with respect to methanol. First-order kinetics versus methanol is typically reported for methanol concentrations lower than 0.1 M and anode potentials lower than 0.5 V versus RHE. The anodic process is typically zero order at methanol concentrations over 0.1 M, indicating that processes (54) + (55) are rate determining. The latter is normally the situation in polymer electrolyte DMFC anodes, which operate on liquid aqueous feed of concentrations typically close to 1 M and at anode potentials in the range 0.35–0.50 V versus RHE [100–102].

Higher coverage by surface oxygen on the PtRu catalyst at higher potentials or ''aging'' of the active surface oxygen species, manifests itself in a strong increase in the Tafel slope typically recorded for the anodic methanol oxidation process at PtRu catalysts above 0.4 V (the Tafel slope typically rises from 120 to 180 mV/decade, or more). Such a high slope apparently indicates an anodically driven surface deactivation process, highlighting the limited potential range within which PtRu has a strong advantage versus Pt as catalyst for methanol oxidation. It is apparently the optimized surface affinity to oxygen within the potential range 0.3–0.5 V versus RHE that makes PtRu a superior methanol oxidation catalyst. The onset of process (54) on PtRu near 0.3 V defines the negative limit of the anode potential range associated with high anodic methanol oxidation activity, whereas excessive surface oxidation at $V > 0.5$ V apparently brings about deactivation of PtRu.

Reference 103 addressed an important challenge of characterizing in situ commercially supplied PtRu anode catalysts, used in polymer electrolyte DMFCs, to enable correlation between the actual state of a PtRu catalyst surface in the anode of an operating DMFC and the anodic activity exhibited. The latter work describes an investigation of three types of unsupported PtRu sample powders supplied from two different commercial sources, which were used to prepare ''inks'' for the fabrication of catalyzed DMFC membranes. The investigation revealed the advantage of reductive preconditioning of such PtRu catalysts and the strong effect that the level of surface oxidation of an as-received PtRu sample could have on the catalyst activity in the DMFC anode. By using CO-stripping voltammetry, it was shown in Ref. 103 that the population of metallic surface sites could be significantly increased by reduction of as-received, unsupported PtRu catalysts of overall surface area ranging between 50 and 80 $m^2\ g^{-1}$. Such a chemical reduction could be achieved by exposure to either hydrogen or to warm aqueous methanol. This finding of enhanced PtRu catalyst activity following reductive treatment is explained by an excess of surface oxide in the as-received PtRu alloy catalyst samples. With increase of metal (= increase in measured CO adsorption sites) on the surface of the PtRu sample, the activity in methanol electrooxidation was shown to increase [103], suggesting that *metal* alloy surface sites are, in high probability, the active sites in the anodic

oxidation of methanol at dispersed, unsupported PtRu catalysts. Various phase oxide species (mainly Ru oxide species) always found in such samples seem to be of no catalytic function beyond a possible role in proton conduction in the catalyst layer according to [103], which was highlighted in Ref. 52.

Whereas the measured CO voltammetric stripping charge indicated the population of metal surface sites under DMFC operation conditions, the stripping peak potential provided some information on PtRu surface alloy composition. From the correlation found between activity and stripping peak potential, it was suggested that metal surfaces closest in composition to $Pt_{0.5}Ru_{0.5}$ apparently enabled most effective stripping of CO and were also the ones most active in anodic methanol oxidation. The latter strong correlation shows that steps (54) + (55) are rate determining for methanol concentrations and anode potentials most relevant in DMFCs [103].

8.3.8.2 The DMFC Membrane: Methanol Crossover and the Impact on the Power/Efficiency Relationship in the DMFC

Starting in the early 1990s, a significant part of the work to develop DMFC technology focused on cells based on polymer electrolytes, primarily Nafion® membranes [3, 4d]. This choice was made on the basis of the attractive properties of membrane electrolyte cells, including the need of only deionized water as the liquid added to the cell (in addition to the fuel itself), good chemical stability of poly(PFSA) membranes, and the favorable chemical environment for the interfacial kinetics of both the anodic and the cathodic process. In the mid-1990s, polymer electrolyte DMFCs had been demonstrated to reach areal power density as high as 0.2 W cm^{-2} in operation on aqueous methanol and air at elevated cell temperatures near 100 °C [104]. This areal maximum power density is about one-third that of a polymer electrolyte hydrogen/air cell and, although obtained with higher catalyst loadings, it demonstrated that the poly(PFSA) electrolyte provides a good basis for DMFCs of high power outputs.

At the same time, however, the use of proton-conducting polymeric membranes for DMFCs presents a significant challenge because of the high permeability of methanol through such membranes. The same problem is not encountered in hydrogen/air cells because of the much lower solubilities of H_2 and O_2 in the hydrated membrane material. The magnitude of the problem is readily understood when comparing the rate of methanol permeation by diffusion through a poly(PFSA) membrane to the DMFC cell current generated at the same temperature. Taking as example a Nafion® 117 membrane at 30 °C, the current density equivalent of methanol permeability under a concentration gradient of 1 M between the fuel and the air side of the membrane is 50 mA cm^{-2} (compared with $\leq$1 mA cm^{-2} for H_2 or O_2), whereas the maximum current density obtained from a DMFC with Nafion 117 membrane electrolyte at the same temperature is around 100 mA cm^{-2}. This means that, unless the rate of crossover is strongly reduced, the cell will be associated with a severe ''chemical short'', as a significant fraction of the fuel provided for the anodic process diffuses instead through the membrane and directly recombes with oxygen at the cathode catalyst. Consequently, unless the concentration gradient driving methanol crossing over the membrane is lowered well below $\Delta C_{MeOH} = 1$ M, fuel utilization in such a cell could be as low as 60% at cell current of 100 mA cm^{-2} and

further lower at lower cell current. Fuel utilization or fuel efficiency is generally calculated as follows:

$$\eta_{\text{fuel}} = \text{rate of anodic conversion of fuel to current/rate of fuel feed to the anode} \quad (57)$$

which, in the case of some current equivalent crossover rate of $J_{\text{crossover}}$, can be written as follows:

$$\eta_{\text{fuel}} = J_{\text{cell}}/[J_{\text{cell}} + J_{\text{crossover}}(J_{\text{cell}})] \quad (58)$$

where $J_{\text{crossover}}$ is recognized in (58) to depend on cell current density, J_{cell}, as discussed in the following text.

Possible approaches to correct for the high permeability of methanol through the poly(PFSA) membrane are as follows:

- Replace the PFSA membrane by another polymeric membrane of lower methanol permeability (while maintaining good cell performance)
- Lower the methanol concentration in the anode feed stream
- Induce strong lowering of the crossover rate with increase in cell current through proper choice of the anode backing layer

A significant number of efforts to prepare membranes of lower methanol permeability for use in DMFCs have been reported. The general types of new DMFC membrane preparations tried to date include (1) replacing poly(PFSA) by sufonated polyaromatic polymers [105], (2) blending polymers of low alcohol permeability (PVDF, for example) with either PFSA or poly(styrene sulfonate) [112], and (3) blending dispersed inorganic material, mainly electronically insulating oxide powders, into a polymer membrane host [106]. The idea behind option (1) is that sulfonated polyaromatics have a lesser tendency for phase separation between hydrophilic and hydrophobic nanodomains of the polymer. It is the very strong tendency of PFSA polymers for nanophase separation, originating from the highly hydrophobic nature of the perfluorocarbon backbone, that readily generates an aqueous, nanopore network of good connectivity in the polymer. It is this well-connected aqueous network that provides high protonic conductivity in PFSA polymers of EWs 1100 or below and also high methanol permeability. In polyaromatic membranes, the hydrophobic character of the aromatic chain is less extreme, resulting in a lesser tendency to form well-connected aqueous nanonetworks at some given water content in the polymer [107]. Table 1 gives an example of the type of proton-conducting polyaromatic membranes synthesized and tested [105].

The approaches, based on the blending of either a polymer of lower intrinsic alcohol permeability [104] or a dispersed oxide into a sulfonated polymer host [106], target lowering methanol permeability by lowering effective porosity and increasing effective tortuosity of the active (aqueous) subphase forming in such composites under the porosity/tortuosity of the unmodified host membrane. In all such attempts to lower methanol permeability, the challenge is not to lower too much of the protonic conductivity of the unmodified membrane in the process. This is a serious challenge because, in all hydrated ionomeric membranes, very similar properties enhance both protonic conductivity and methanol permeability, including the overall water content and the nature of a separated aqueous network of nanopores that form at some level of hydration. Consequently, as the properties of aqueous nanopore networks are changed by any

of the approaches described in Refs 104, 106, and 107, the fractional change in methanol and proton transport rates tend to be similar. Exceptions, where significant selectivity enhancement was reported for proton over methanol transport compared with the poly(PFSA) membrane, are few and in-between [104, 108].

The ratio $J_{cell}/J_{crossover}$ is expected to increase with cell temperature because the dependence of J_{cell} on temperature is determined by a relatively high activation energy of oxygen and methanol electrode processes, roughly corresponding to doubling of the interfacial rate in the relevant temperature range per raise of T_{cell} by 20 °C. Methanol crossover is, on the other hand, an aqueous (sub)phase transport process associated with a weaker rate increase of less than 50% per raise of T_{cell} by 20 °C. However, raising cell temperature is a limited option in most portable power applications of DMFCs.

One possible approach to lower the rate of fuel crossover for a given membrane is to lower the methanol concentration in the anode feed stream. Lowering of the feed concentration is a viable option, in principle, because the overall process of anodic methanol oxidation is zero order in methanol concentration, at least down to 0.1 M. Operation with very low concentrations of recirculated aqueous methanol in the anode compartment would, however, require high rates of anode feed recirculation and some control of methanol concentration in the recirculated anode feed stream. The more effective option of lowering the rate of methanol crossover in an operating cell is to rely on the cell current as the means for lowering the concentration of methanol on the anode side of the membrane. This can be achieved by impeding methanol transport though the anode backing, such that the limiting current of methanol through the backing is only about 10% higher than the cell current at the design cell voltage [109]. Under such conditions, as described schematically in Fig. 48, the methanol concentration in the membrane next to the anode catalyst is a small fraction of the methanol concentration in the anode flow stream (i.e., on the opposite side of the anode backing) and, as a result, the methanol crossover rate falls significantly versus the crossover rate measured in the same cell at zero cell current. To quantify the relationship between $J_{crossover}$ and J_{cell} in a DMFC targeting such large methanol concentration drops across the anode backing under cell current, it is first recognized that, under open-circuit conditions, the total impedance to methanol flow is a series combination of the anode backing and membrane flow impedances. $J_{crossover\ @\ J_{cell}=0}$ can therefore be described as follows:

$$1/J_{crossover\ @\ J_{cell}=0} = 1/J^*_{mem} + 1/J^*_{an,back.} \quad (59)$$

where J^*_{mem} and $J^*_{an,back}$ are respectively limiting current densities for methanol transport through the membrane and through the anode backing under a concentration gradient equal to the methanol concentration at the anode backing/flow channel interface (Eq. 59 assumes a small concentration drop across the catalyst layer itself). When the cell is under current J_{cell}, $J_{crossover}$ is given by the difference between the current equivalent of the flux of methanol across the anode backing layer and the cell (= anode) current, that is

$$J_{crossover} = J_{an,back} - J_{cell} \quad (60)$$

From the last two equations, the relationship between the rate of crossover from the

anode flow field to the cathode of a DMFC can be derived as follows:

$$J_{\text{crossover}}/J_{\text{crossover @ } J_{\text{cell}}=0} = 1 - (J_{\text{cell}}/J^{*}_{\text{an,back}}) \quad (61)$$

The last equation reveals that, to the degree that a DMFC can be operated close to the anode-limiting current (corresponding to the concentration of the feed stream), the crossover rate in the cell under current can be very low, even in cases of backing and membrane combinations that allow high rates of crossover at open circuit. With fuel utilization in the DMFC given, in general, by $\eta_{\text{fuel}} = J_{\text{cell}}/(J_{\text{cell}} + J_{\text{crossover}})$, the relationship $J_{\text{crossover}}$ (J_{cell}) derived above (Eq. (61) allows to calculate fuel utilization at some J_{cell} in a DMFC of the given membrane and anode backing permeability of methanol:

$$\eta_{\text{fuel}} = J_{\text{cell}}/\{J_{\text{cell}} + J_{\text{crossover @ } J_{\text{cell}}=0}(1 - J_{\text{cell}}/J^{*}_{\text{an,back}})\} \quad (62)$$

This expression shows that fuel utilization levels as high as 90% can be achieved even when the membrane is quite ''leaky'' to fuel, that is, when $J_{\text{crossover @ } J_{\text{cell}}=0}$ is quite large. For example, with $J_{\text{crossover @ } J_{\text{cell}}=0}$ as high as 50% of the cell current density targeted, operation at $J_{\text{cell}}/J_{\text{lim,an}} = 0.85$ enables, according to Eq. (62), η_{fuel} as high as 93% with the ''leaky'' membrane.

A special case of Eq. (62) arises when the DMFC anode is (methanol) dead-ended, that is, all methanol fed into the anode at a current equivalent rate of J_{feed} either reacts at the anode or recombines with oxygen at the cathode. These operation conditions apply in 1–3 W, passive DMFC platforms developed for consumer electronics applications described later in this section. Under these conditions, $\eta_{\text{fuel}} = J_{\text{cell}}/J_{\text{fuel feed}}$ and when the cell limiting current is determined by $J_{\text{lim,an}}$,

$$\eta_{\text{fuel}} = J_{\text{cell}}/J_{\text{fuel feed}} = J_{\text{cell}}/J_{\text{lim,an}} \quad (63)$$

that is, fuel efficiency is directly proportional to cell current. It is also clear from Eq. (63) that in operation at a $J_{\text{cell}}/J_{\text{lim,an}}$ ratio near unity, the fuel efficiency in the DMFC is fully determined by the match between cell current and fuel feed rate, irrespective of the degree of methanol permeability through the membrane.

From (63), it is clear that, to achieve high η_{fuel} at some given design current and cell temperature, the anode-limiting current has to be set just slightly above the target current such that the ratio $J_{\text{cell}}/J_{\text{lim,back.}}$ is maintained slightly under unity. Consequently, if high fuel efficiency is to be achieved in cell operation through a significant dynamic range, the anode feed concentration has to be adjusted proportionately to the variations in cell current demand. This is possible to achieve within limitations of the system's response time to required feed rate variations, as would be defined by the design of the fuel-delivery subsystem.

Another interesting challenge to operation at 90% the anode-limiting current as a means to achieve high fuel utilization is the magnitude of cell voltage losses incurred in operation so close to the anode mass-transport limit. Additional cell voltage losses caused by strong transport limitations could, at least in principle, overweigh the fuel utilization gains. Such a suspicion arises when examining the current voltage dependence ordinarily written for mixed cell current control by interfacial kinetics and mass transport: an extra loss of cell voltage, $\delta\eta_{\text{an}} = b_{\text{an}} \log\{1/(1 - J_{\text{cell}}/J_{\text{lim,an}})\}$ would be the expectation according to Eq. (10). At 90%

the limiting current, this corresponds to an extra anode voltage loss numerically equal to b_{an}, that is, $\geq$100 mV. If this were indeed the case, working near the anode-limiting current would not be a very viable option. In reality, however, such mass transport losses are not incurred because they are projected for a first-order interfacial process, whereas the anodic process in the DMFC is zero order in methanol (see Eq. 54, 55). The polarization curve for a DMFC with an air cathode (first order in oxygen) and a methanol anode (zero order in methanol) is given for an ultrathin electrode as follows:

For $J < J_{critical}$:

$$V_{cell} = V_0 - b_{an}\log_{10}\left(\frac{J}{a_{an}J_{0,an}}\right) - b_{cath}\log_{10}\left[\frac{\left(\frac{J}{a_{cath}J_{0,cath}}\right)}{\left(1-\frac{J}{J_{lim,cath}}\right)}\right] - JR_{ohm}$$

For $J > J_{critical}$:

$$V_{cell} = V_0 - b_{an}\log_{10}\left[\left(\frac{J}{a_{an}J_{0,an}}\right)\frac{\left(1-\frac{J_{crit}}{J_{lim,an}}\right)}{\left(1-\frac{J}{J_{lim,an}}\right)}\right] - b_{cath}\log_{10}\left[\frac{\left(\frac{J}{a_{cath}J_{0,cath}}\right)}{\left(1-\frac{J}{J_{lim,cath}}\right)}\right] - JR_{ohm} \quad (64)$$

where J_{crit} is the cell current density at which the concentration of methanol at the catalyst surface had dropped sufficiently to switch from zero order (step 54 rate limiting) to first order (step 52 becoming rate limiting). Up to J_{crit}, the anode does not suffer any mass-transport loss because the rate of the interfacial process does not drop as the local concentration of methanol at the catalyst surface keeps falling.

As shown in Fig. 52(a), the polarization curves calculated for the DMFC according to Eq. (64) exhibit a sharp fall in cell voltage, occurring very close to the anode-limiting current. The sharpness of the fall increases the higher J_{crit}, that is, when zero-order characteristics prevail at cell current very close to J_{lim}. The simulation in Fig. 52(a) is assumed to be $J_{crit} = J_{lim,an} - 0.01$ A cm^{-2}. Figure 52(a) displays three calculated polarization curves, corresponding to 3 orders of magnitude of the exchange current density of the cathode and anode, assumed in this simulation to have identical values (this assumption is quite reasonable since in both cases measurable currents are observed first at a similar overpotential of 0.3 V). The "base case" curve reflects the state-of-the-art for DMFCs operating slightly above ambient temperature and the other two curves demonstrate the strong DMFC performance advances expected from raising the rates of both electrocatalytic processes.

Before the next breakthrough in electrocatalysis of the methanol anode and/or the air cathode, the only viable way to increase the exchange current density by an order of magnitude (without further increase in the relatively high catalyst loading required in the DMFC) is to increase the cell temperature. In fact, the lower and mid curves in Fig. 52(a) describe well a DMFC operating close to ambient conditions and near 80 °C, respectively. Figure 53 describes a measured DMFC polarization curve, recorded at 80 °C with liquid aqueous 0.5 M methanol as anode feed and air supplied near ambient pressure, showing a power output near 100 mW cm^{-2} at 0.5 V – close to four times higher than generated at the same cell voltage at a cell temperature near ambient.

Figure 52(b) describes the variations of cell power output and cell energy–conversion efficiency with cell current in a DMFC with fuel utilization is calculated as $J_{cell}/J_{lim,an}$ (Eq. 63). This expression corresponds to the specific case of a (methanol) dead-ended anode. Unlike the case of the hydrogen/air cell, where fuel efficiency is high at any cell current, the power output and conversion efficiency of the DMFC increase hand in hand with rise in cell current. Both power output and conversion efficiency reach their peak slightly short of the cell (anode) limiting current and both drop with further increase in cell current. In fact, since $\eta_{cell} = \eta_{fuel} \times \eta_v$, η_{fuel} given in the case of a dead-ended methanol anode as $J_{cell}/J_{lim,an}$ and η_v is given by V_{cell}/V^0_{cell}, both η_{cell} and P_{cell} and increase with the product $J_{cell}V_{cell}$ and, consequently, peak power and peak efficiency are reached at an identical point on the polarization curve (Fig. 52b). A very similar pattern applies also for the case of a DMFC with a recirculated anode feed stream of controlled methanol concentration, as can be seen from Eq. (62), which is applicable in the latter case.

To summarize this description of the "leaky membrane challenge" in the DMFC and how it can be best addressed, the choice to date has been to either (1) develop an alternative membrane that will exhibit lower methanol permeability while maintaining high protonic conductivity and well-performing MEA or (2) stay with poly(PFSA) membranes and MEA technology and address the leaky nature of such membranes in operation with methanol by lowering the methanol concentration next on the anode side of the membrane in a cell under load. The latter is achieved by (1) lowering the concentration in the anode flow channel and (2) matching the anode permeability to the cell current demand to further significantly lower the methanol concentration at the membrane surface under current (Fig. 51). The alternative membrane route involves a significant materials R&D challenge: achieving high proton mobility-to-methanol mobility ratio in a well-conducting protonic membrane. In contrast, the optimized components and system route, uses other cell components and system controls to address the leaky nature of poly(PFSA) membranes in DMFCs. As in other similar cases, it is easier to implement earlier on a solution based on a proven and tested membrane material while applying proper correction(s) for its deficiencies at the stack and system design and control levels.

8.3.8.3 The DMFC Air Cathode

The air cathode in the DMFC presents some specific challenges beyond those documented for the hydrogen/air PEFC (see section 8.3.7.2). These challenges are related to the possible flux into the cathode

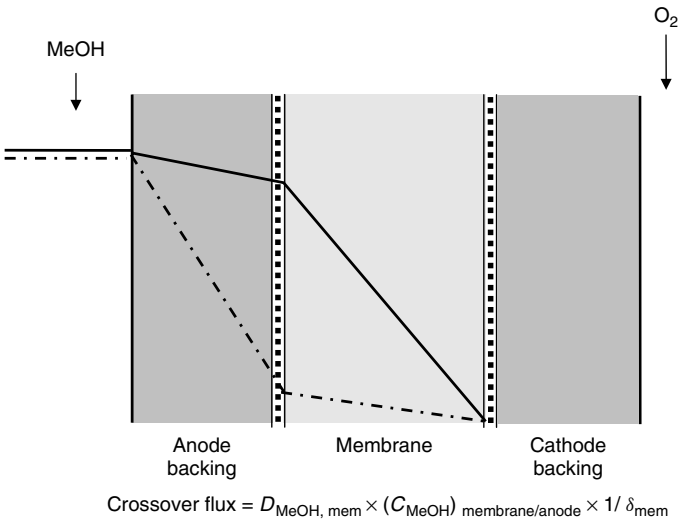


Fig. 51 Lowering the rate of methanol crossover by operation at cell current close to the limiting current defined by the anode backing layer. Solid – MeOH concentration profile at open circuit; dashed – MeOH concentration profile near $J_{lim,an}$.

of methanol and water from a DMFC anode fed by liquid aqueous methanol. This liquid flux across the membrane is driven primarily by a concentration gradient in the case of methanol and by electroosmotic drag, that is, cell current in the case of water. Possible impacts of methanol penetration into the cathode start from the simplest case [110], where the processes of oxygen reduction and methanol oxidation each occur without interference by the other and the component cathode currents of opposite signs add algebraically. A more severe situation at the DMFC cathode corresponds to cathode ''flooding'', which can occur as a result of loss of hydrophobic nature of the cathode backing following wetting by aqueous methanol.

Noninteracting processes of oxygen reduction and methanol oxidation at the DMFC cathode result in a ''mixed potential'' corresponding to increase of cathode overpotential at some given cell current. Under most common DMFC operation conditions, the oxygen reduction process is under interfacial kinetics control and the methanol oxidation process is determined by the methanol flux across the membrane, $J_{crossover}$, (Eqs 59–61). The increase in cathode overpotential is therefore be given by

$$\begin{aligned}\delta\eta_{cath} &= b_{cath}\log[J^*_{cell}/(J^*_{cell} - J_{crossover})] \\ &= b_{cath}\log[J^*_{cell}/J_{cell})] \end{aligned} \quad (65)$$

where J^*_{cell} is the oxygen reduction current component at the cathode potential corresponding to a measured cell current density of $J_{cell} = J^*_{cell} - J_{crossover}$. From Eq. (65), it is clear that the effect of some given flux, $J_{crossover}$, of methanol penetrating into the cathode on η_{cath}, depends on the ratio $J_{crossover}/J^*_{cell}$, causing cathode potential shifts which are largest at the smallest cell current densities (higher V_{cell}), becoming negligible when J^*_{cell} is

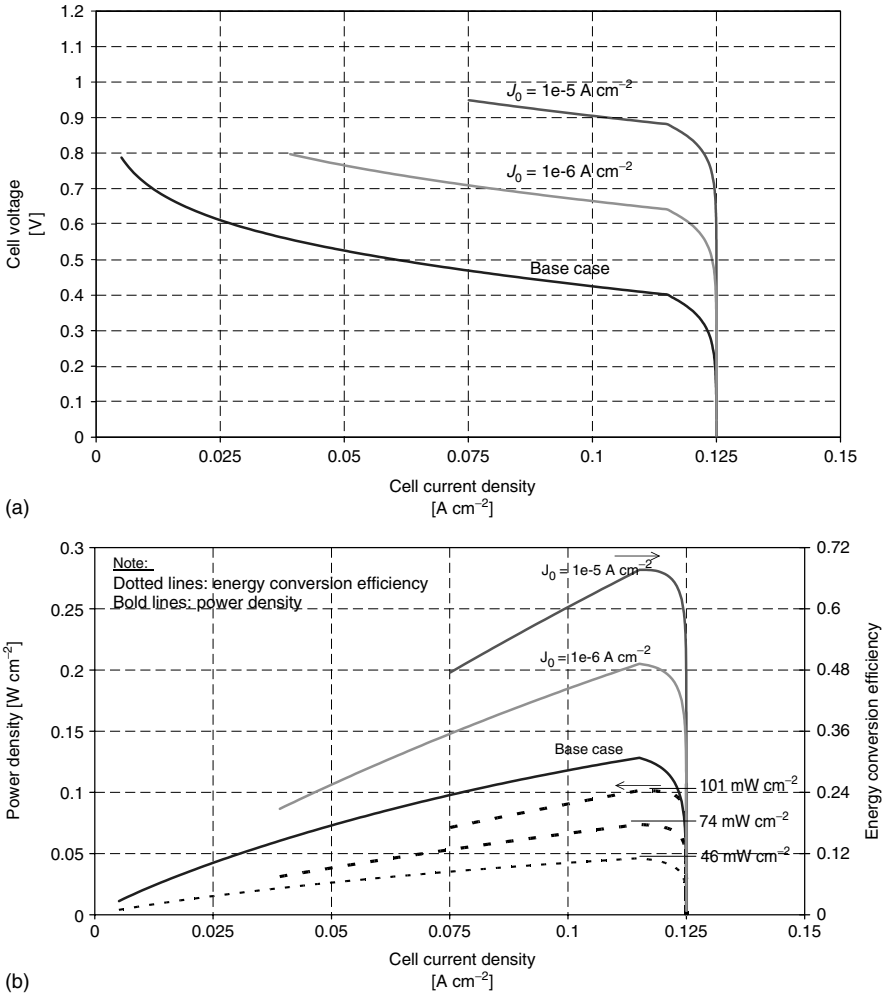


Fig. 52 (a) DMFC polarization curves calculated according to Eq. (64) which assumes ultrathin electrodes, negligible effect of methanol crossover, and zero-order anode process in methanol up to $J_{crit} = J_{lim,an} - 0.01$ A cm^{-2}. The three calculated polarization curves, corresponding to 3 orders of magnitude of the exchange current density of the cathode and anode (assumed identical), reflect state-of-the-art polymer electrolyte DMFC performance near ambient conditions (bottom curve) and near 80 °C (mid curve), as well as the further improvement in cell performance expected at 80 °C if/when an increase by another order of magnitude of the catalytic activities of DMFC cathode and anode is achieved; (b) calculated power output and energy-conversion efficiency as function of cell polarization, corresponding to the three cases in (a).
The specific parameters used for calculating the curves in Fig. 52 are as follows:

$$J_{0,an} = J_{0,cath} = 10^{-7}\ \text{A cm}^{-2}, \qquad a_{an} = a_{cath} = 1000\ \text{cm}^2\ \text{of Pt/cm}^2\ \text{of membrane area},$$
$$b_{an} = b_{cath} = 120\ \text{mV/decade of current density}, \qquad J_{lim,an} = 0.125\ \text{A cm}^{-2},$$
$$J_{lim,cath} = 0.35\ \text{A cm}^{-2}, \qquad R_{ohm} = 0.375\ \Omega\ \text{cm}^2, \qquad J_{crit} = J_{lim,an} - 0.01\ \text{A cm}^{-2}.$$

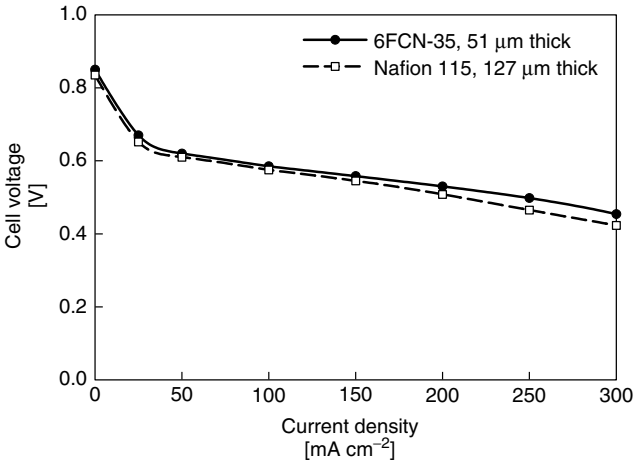


Fig. 53 Comparison of DMFC polarizations curves at 80 °C with anode feed of 0.5 M methanol solution and the air cathode operating at zero backpressure for a cell with a Nafion 115 membrane and a cell with a polyaromatic membrane of lower methanol permeability [105].

significantly larger than $J_{\text{crossover}}$. From Eq. (65), the cathode loss associated with crossover drops to 8 mV when the ratio $J_{\text{crossover}}/J^{*}_{\text{cell}}$ drops to 0.25, corresponding to a conversion efficiency loss of two percentage points versus the case of zero methanol crossover. As explained in detail in the previous section, operation near the cell (anode)–limiting current enables strong reduction in $J_{\text{crossover}}$ and, consequently, significant reduction in this type of DMFC cathode voltage loss.

Cathode voltage losses incurred by methanol crossover could, however, go well beyond the loss calculated for such a "noninteraction" case. Cathode losses could be much more substantial when a film of aqueous methanol builds up in the cathode under some DMFC operation conditions and blocks oxygen access to the cathode catalyst. The resulting flooding of the cathode is typically a stronger challenge in DMFCs with aqueous methanol feed streams than it is in hydrogen/air PEFCs because of the high rates of liquid penetration from a liquid-fed DMFC anode across the membrane into the cathode. The water flux is driven by an electroosmotic drag of 2.5–3.5 water molecules per proton, that is, nearly 20 molecules of water per molecule of methanol reacting at the anode travel across the membrane in a DMFC with an aqueous methanol anode feed [33]. Methanol carried along into the cathode further exacerbates cathode flooding by negatively affecting the cathode wetproofing characteristics. Water buildup in the DMFC cathode can, in fact, be reduced by minimizing the rate of methanol crossover through the range of DMFC operation conditions, that is, adjusting the fuel feed rate to the cell current as the latter varies (see Eq. 61). Additionally, proper design of the cathode flow field, to achieve effective liquid removal by airflow across the cathode surface can assist in sweeping excess cathode liquid with the air exhaust. Cathode features determining the

efficiency of water removal include a cathode backing with a proper degree of openness and strong hydrophobic surface characteristics and, at the same time, flow channels with hydrophilic surfaces, facilitating transfer of water droplets from the cathode GDL to the airflow channel (see Fig. 47). Figure 54 shows that, with effective design of the DMFC air cathode, cell performance is maintained high at very low stoichiometric flow of air. In fact, it is within 90% of the performance measured with very high airflow rates, shown in Fig. 53. The ability to operate at such low airflow rates has clear advantages in both minimizing parasitic power losses and in avoiding excessive water losses from the cell.

8.3.9 Potential Applications of DMFCs in Portable Power Sources

Significant interest has been shown recently by consumer electronics manufacturers and by the military sector in low-temperature portable fuel cells, particularly portable DMFC. Both civilian and military sectors are interested in a next-generation, small portable power source of energy density that significantly exceeds that of leading-edge battery technologies. This means a power system energy density well beyond 400 Wh L^{-1} (0.4 Wh cm^{-3}) and/or a specific energy well beyond 200 Wh kg^{-1} (0.2 Wh g^{-1}) to be achieved while providing the power for the full dynamic range of the load in a given device. Advanced power sources able to meet and beat these specifications will allow greatly extended run times per recharge (refuel) for mobile phones, laptop computers, and digital cameras – and also for radios, night-vision goggles, global positioning systems, and laser-targeting units in the military arena. Perhaps most even more significantly, higher-energy-density power sources are likely to be essential building blocks in "converged" handheld communications devices that combine

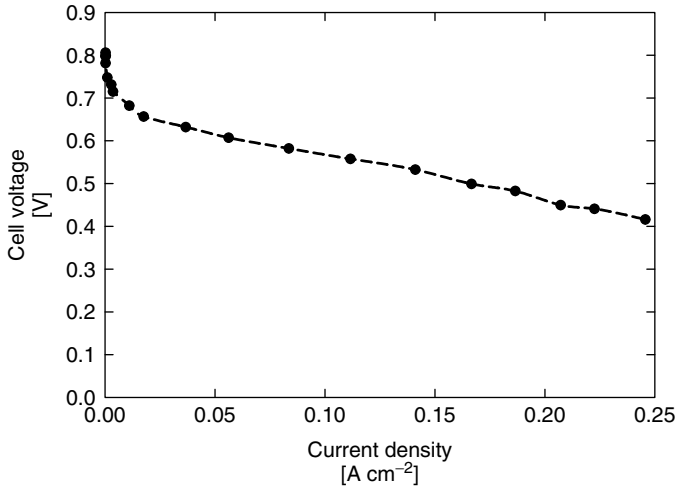


Fig. 54 DMFC polarization curve (80 °C) for a 50-cm^2 cell using a Nafion 117 membrane, with anode feed of 0.5 M methanol solution and the air cathode operating at zero backpressure and limited stoichiometric flow of 1.9 at 0.5 V.

color-display technology with voice, broadband data, and video functionality – all in a single "power-hungry" platform.

The strong potential advantage of fuel-cell power sources over batteries originates from the high energy density of the fuel. For low-temperature fuel cells ($T_{cell} < 100\,^{\circ}C$) applicable for man-portable devices, the theoretical specific energies of the two most suitable fuels are 33 Wh g^{-1} for hydrogen and 6.1 Wh g^{-1} for methanol. In comparison, advanced rechargeable lithium-ion batteries have a specific energy of around 0.2 Wh g^{-1} – less than 1% of the energy density of neat hydrogen and 3% that of neat methanol. Such a comparison is obviously skewed, because the battery comprises a complete power source in one package, whereas the fuel for a fuel cell still has to be packaged with the chemical-to-electrical energy-conversion device (the fuel cell itself) plus any further auxiliaries, including power conditioning and power system controls. Figure 55 is a building block scheme of a portable power source based on an air-breathing fuel cell fueled by liquid methanol. The power system includes a fuel-containment chamber and the volume fraction taken by the fuel is shown to be relatively large, as this is prerequisite for overall system energy density advantage over the incumbent battery technology. The total volume of portable power sources of this type, designed for laptops, for example, will have to be limited to about 200 cm^3, whereas for handsets, it will have to ultimately drop under 30 cm^3. These very small volumes will have to accommodate enough methanol fuel to provide significantly more energy to the load than an incumbent battery of similar volume. The methanol fuel will have to share, however, the small total volume available, with the fuel cell itself and all auxiliaries required to generate power outputs in the range of 10–30 W for a laptop computer and 1–3 W for a handset.

To summarize the effects of these key parameters, if a fraction f_{fuel}^{wt} of the total weight of a power system is the fuel itself and the system's overall energy-conversion efficiency (chemical-to-electrical energy to the load) is η_{system}, then the actual specific energy of the power pack, E_{system}^{wt} (Wh g^{-1}), will be given by Eq. (1).

$$E_{system}^{wt} = f_{fuel}^{wt} \times E_{fuel}^{wt} \times \eta_{system} \quad (66)$$

where f_{fuel}^{wt} is the weight fraction of fuel in the system, E_{fuel}^{wt} is the theoretical specific energy of the fuel stored, and η_{system}

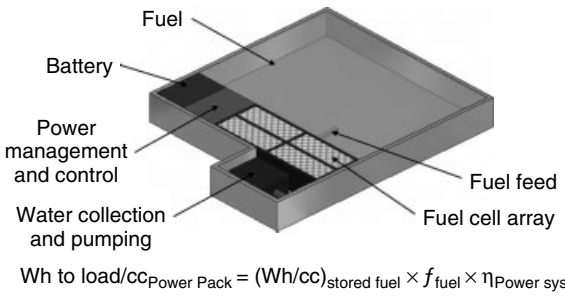


Fig. 55 Building blocks scheme for a micro-DMFC Power Pack and the resulting impact on the system energy density, that is, the energy delivered to the load per unit volume of the power pack.

is system efficiency for conversion of the chemical energy of the fuel to electrical energy to the load. For example, for a fuel-cell power system in which methanol fuel accounts for 50% of the weight and the system conversion efficiency is 25%, E^{wt}_{system} is calculated from Eq. (1) as 0.75 Wh g^{-1}. In other words, the specific energy advantage over a lithium-ion battery of specific energy of 0.2 Wh g^{-1} is 3.75:1. The fuel-cell potential power-pack advantage is even stronger when it comes to carrying or supplying additional energy for extended use, that is, when the specific energy of the replacement fuel cartridge becomes a relevant parameter ($f^{wt}_{fuel} = 1$ in Eq. 66). In this case, the specific energy advantage over a replacement lithium battery is expected at 7.5:1. This significantly higher ratio (and stronger value proposition) is relevant for cases of long use without the ability to recharge – for example, in military applications, when longer use time requires supply of additional, fully charged batteries, whereas for a fuel cell power system, additional fuel cartridges. It also comes into play in the context of a ''perfectly wireless world'', in which lightweight replacement fuel cartridges could be used to eliminate the need for cord connections between handheld devices and a recharging power source. In the example given above, carrying replacement methanol cartridges for a cord-independent, fuel-cell-powered device is 7.5 times more mass-effective versus carrying replacement batteries.

For compact portable consumer electronics applications like mobile phones, it is often the volume rather than the weight of the power source that is the key figure of merit in terms of system packaging. Under these circumstances, it is the energy density, rather than the specific energy, that is of greater interest. The equivalent of Eq. (1) will then be Eq. (2).

$$E^{v}_{system} = f^{v}_{fuel} \times E^{v}_{fuel} \times \eta_{system} \quad (67)$$

Here, f^{v}_{fuel} is the fuel-volume fraction in the system, E^{v}_{fuel} is the theoretical energy density of the fuel, expressed, for example, in Wh cm^{-3}; η_{system} is, as before, the system conversion efficiency for chemical energy of fuel to electrical energy to the load and E^{v}_{system} is the effective energy density of the complete power system in Wh cm^{-3}.

From the discussion above, it is clear that designers of small-scale fuel-cell-based power sources must adhere to the following priorities:

1. Maximize E^{wt}_{fuel} (Eq. 66) using the following guidelines:
 (i) Realistic fuel storage options, based on either neat, energy-rich fuel or on a fuel storage form maintaining a large volume/weight fraction of a neat, energy-rich fuel.
 (ii) Use of fuel types acceptable for man-portable applications, considering safety requirements in such applications.
2. Maximize f^{wt}_{fuel} (Eq. 66) by minimizing $(1 - f^{wt}_{fuel})$, the fraction of the power-pack weight (or volume) occupied by energy-conversion components of no energy content – that is, the fuel cell itself plus any auxiliaries The latter necessitates the following in turn:
 (i) Fuel-cell stack technology of minimum weight (or volume) requirement to generate the average power demand of a portable device under the benign operating conditions dictated by handheld applications.

(ii) A simple overall power system to minimize auxiliaries weight/volume and to increase overall power system reliability.

3. Maximize η_{system} by raising (Eq. 7) the following:

 (i) Fuel utilization
 (ii) The fuel-cell voltage at the power level demanded by the load

At this point in time, the choice of fuel for portable power applications is effectively limited to hydrogen and methanol. Proper storage of hydrogen for portable fuel-cell applications is challenging, however, and the most promising solution appears to be the use of metal hydrides when small volume is of higher premium than lower weight, or the use of ''chemical hydrides'' in the reverse case. In the former case, heating the metal hydride releases free hydrogen as fuel feed to the cell, while the latter typically comprises a hydrogen-releasing chemical reaction of a solid like sodium borohydride, which undergoes a controllable reaction with water over a proper catalyst. A variant of the latter has also been proposed: direct electrochemical oxidation of a soluble hydride – potassium borohydride fuel dissolved in an aqueous alkaline solution. The overall weight percent of hydrogen in relevant metal hydrides (those releasing hydrogen at temperatures under 100 °C) is of the order of 1–2%. For the chemical hydrides, assuming the need to carry the water for the reaction, this figure can reach 6–8%. The case of a concentrated aqueous solution of sodium borohydride, discussed in the following text, corresponds to theoretical energy density of 1.8 Wh g^{-1} – about 40% the theoretical energy density of liquid methanol which stores energy at 6.1 Wh g^{-1}. The advantage of hydrogen fuel, though, is its superior electrochemical activity (see section 8.3.7.11), which is of particular significance at low fuel-cell temperatures, facilitating the achievement of higher stack power densities and thereby leaving more precious system space for fuel.This is of high premium in applications with the lowest volume availability, for example in potential applications in powering cell phones. The choice between the two candidate fuels, hydrogenated nd methanol, for portable power applications must be made with all the above in mind, in additions to other important features including fuel distribution infrastructure and power system safety can be readily understood, the choice may be a strong function of a specific application, particularly the target use time per single refuel, the market sector being consumer electronics or the military.

8.3.9.1 Microfuel Cell System Engineering Challenges

The two fundamental prerequisites for maximizing the power density and energy-conversion efficiency in miniaturized fuel cells in general and miniaturized DMFCs, in particular, are (1) sufficient power and energy to the load per unit area of the electrodes to be generated at cell inner temperatures as low as 45–60 °C and (2) minimal overall volume/weight of the non-fuel-containing fraction of the power system. With optimized PtRu alloy anode catalysts and well-prepared electrode structures, DMFCs can generate, at cell temperatures in the range of 45–60 °C, an areal power density of 40–60 mW cm^{-2} at cell voltage $\geq$ 0.4 V. This can be achieved with no active air feed to the cathode, such that the cell is operating in an ''air-breathing'' mode. Assuming a typical unit cell thickness of 3–5 mm, a DMFC stack volume

power density of 100 mW cm^{-3} under cell conversion efficiency near 30% should be possible under these benign operation conditions relevant to man-portable power. While this power density is lower by factor 3–5 versus that of a hydrogen–air fuel cell operating at a similar cell temperature. However, the advantages of methanol as liquid fuel of high energy density, together with the potential for high simplicity of a DMFC system can provide a substantial advantage to the DMFC option, particularly so when the target use time per refuel is longer than 48 hours.

The plots in Fig. 56 reveal the potential of a 1-W DMFC system for lowering the system volume versus that of a lithium-ion battery of energy density 0.2 Wh cm^{-3}. The volumes are compared for a range of overall system energy content demands or overall use time (the number of Whs equals the number of hours of use at the 1-W power output considered). At fuel-cell stack power density of 100 mW cm^{-3} (100 W L^{-1}), the energy density of a DMFC system is projected to exceed that of the lithium-ion battery for use times exceeding 4–6 hours – assuming total auxiliaries volume of 10 cm^3, equal to that of the fuel cell stack itself and assuming system conversion efficiency near 30%. The additional parallel lines in the figure show that, even if the volume

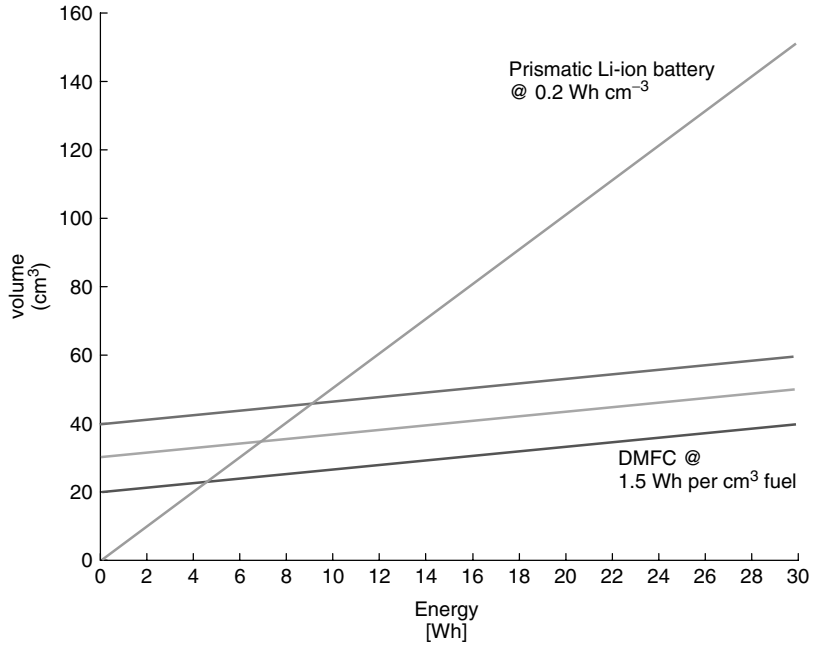


Fig. 56 System volume for a 1-W DMFC power source versus number of watt-hours stored (= hours of use time at 1-W to the load per single fuel cartridge) compared to the volume of a prismatic Li-ion battery of the same number of watt-hours. The three cases shown are for three different volumes of the nonfuel-containing components of the DMFC system, that is, stack + auxiliaries. Assumptions for the DMFC system are: stack power density of 50–100 W L^{-1} (0.05–0.1 W cm^{-3}), auxiliaries volume = stack volume, and system conversion efficiency = 30%.

of the non-fuel-containing part of the system is ×1.5 or ×2 larger than the best case of 20 cm^3 considered here, significant volume reductions are still possible when slightly longer use times of 7–10 hours without recharge (or refuel). Figure 56 shows that the effective energy content of the fuel, $E^{v}_{fuel} \times \eta_{system}$, is a key parameter in establishing energy-density advantage of the overall fuel-cell system over the Li battery. At the same time, the advantages of higher cell/stack power densities are also obvious from the same figure, allowing a larger space/weight for the fuel on board and enabling energy-density advantages over batteries for relatively shorter use times.

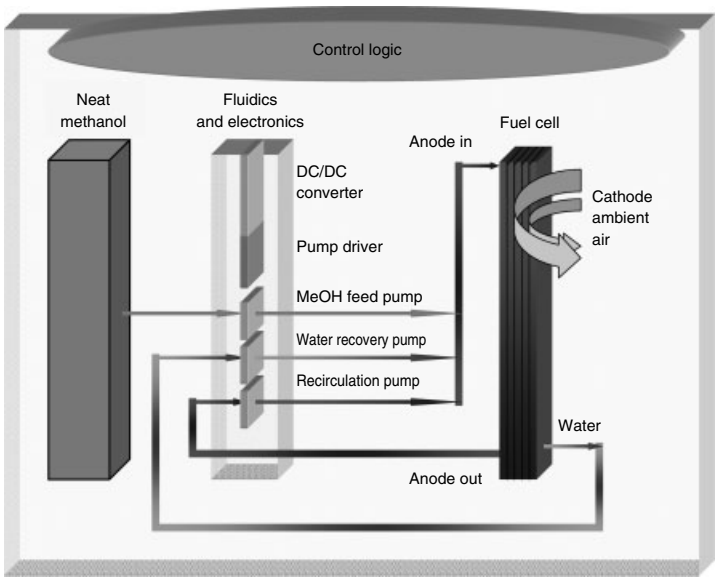
The proper trade-off between higher-energy density of the fuel and higher cell/stack power density is likely to be decided on the basis of the target application, as well as the target use time for the specific application (see Fig. 56). Choices to date to develop a certain type of small-scale fuel-cell power source have been directed by such considerations, as well as by local technical expertise with either hydrogen-fueled PEFCs, direct methanol–fueled DMFCs, or reformed methanol–fed PEFC. Among hydrogen-fueled, small-scale fuel-cell systems described to date, one which has been developed so far into several demonstrated prototypes is based on fuel stored as aqueous alkaline solution of sodium borohydride at 20–25% by weight. The borohydride reacts with the solvent water when the solution is passed over a metal catalyst to release hydrogen according to

$$NaBH_4 + 2H_2O = NaBO_2 + 4H_2 \quad (68)$$

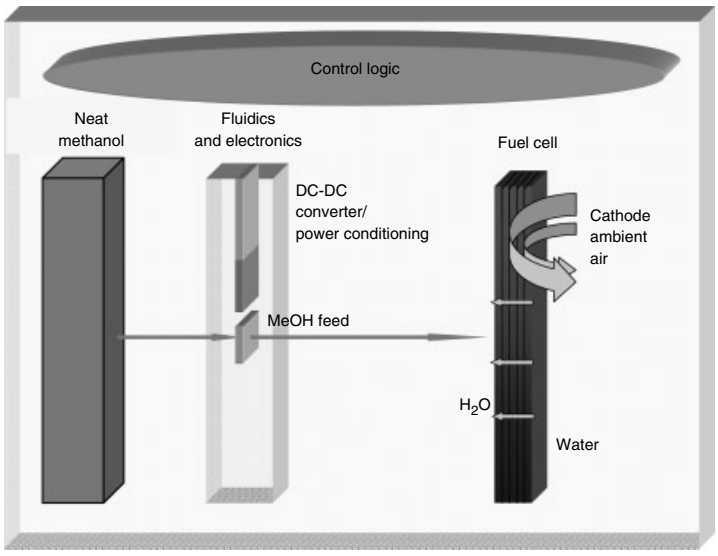
From Eq. (68), at 20 wt%, the aqueous solution of $NaBH_4$ amounts to carrying hydrogen at somewhat above 4 wt%. Consequently, the theoretical energy content of this type of fuel is near 1.3 Wh g^{-1} of fuel and, at 60% conversion efficiency (PEFC operating at $V_{cell} = 0.7$ V), the effective energy content of the aqueous borohydride fuel could reach 0.8 Wh g^{-1} of the fuel solution. A further energy content hike enabled by somewhat higher $NaBH_4$ concentrations in the aqueous hydride fuel has been mentioned as a possibility. In comparison, a significantly higher effective fuel energy density, $E^{wt}_{fuel} \times \eta_{system}$, of 1.4 Wh g^{-1} (1.8 Wh cm^{-3}) of fuel stored, is achievable with a DMFC system storing neat methanol with the fuel cell operating at $V_{cell} = 0.5$ V. Effective utilization of this superior energy density of the fuel to achieve high system energy density requires maximizing of the DMFC stack power density and minimizing of the weight/volume of any auxiliary component. Until very recently, the only DMFC system designs recognized viable were, however, quite complex, largely because of the difficulties associated with water management of water. The challenge has been to reconcile the target of neat (100%) methanol in the fuel tank/cartridge as required to maximize system energy density, with the need for a sufficient water supply to the anode. The latter required for the anode process itself (see Eq. 56) and for dilution of the anode feed stream to lower the rate of methanol crossover by diluting the anode feed. The mainstream approach to this problem has been to provide the water to the anode from the air side of the DMFC, where water is the product of the cathode process using an ''active-system approach'' in which water collected from the cathode is pumped external to the cell and mixes with neat methanol from the fuel tank to provide a dilute aqueous methanol as anode feed. The concept for this recirculated anode loop configuration using active water pumping from

the cathode outlet and into the anode feed stream is shown in Fig. 57(a). Implementation of a complete water-management unit around the DMFC in an active system of this type has been demonstrated by several developers at the 20–40 W power levels. However, as Fig. 57(a) clearly implies, such a DMFC system falls short in terms of the targeted system simplicity, particularly for smaller power sources down at the 1-W level. The collection and pumping of the cathode water product requires a number of moving parts that are involved with extra system volume/weight, parasitic power losses, and system complexities – all translating into oversize and reliability issues for DMFC power sources of 1–3 W.

One way to achieve a simpler DMFC system, albeit at significant expense of energy content would be to include water as a component of the liquid in the fuel cartridge, thereby eliminating the need to collect water and feed it back to the anode in the operating system. However, the energy content penalty associated with such an approach would be prohibitive: any sacrifice of neat fuel weight in a DMFC system significantly impacts the competitive edge versus incumbent battery technology. (From Fig. 56, consider the effect of giving up, e.g., 50% of E^{v}_{fuel} by carrying 50% aqueous methanol in the cartridge.) A recent novel approach to the water-management challenge in DMFCs has been developed and implemented at prototype level by MTI MicroFuel Cells (United States). In the latter approach, two key innovations are: (1) an effective mode of 100% methanol fuel delivery directly to the anode and (2) of the required flux spontaneous generation of water *within the cell* from the air cathode back to the methanol anode (Fig. 57b). The fuel-delivery mode developed enables fine-tuning of the overall flux of neat methanol from reservoir to the anode. This controlled flux of 100% methanol mixes at the anode with water supplied from the cathode internally by a spontaneously occurring flux of water from cathode to anode across the thickness dimension of the cell. This mode of operation eliminates any water collection and pumping in a cell fed directly by 100% methanol fuel, resulting in superior DMFC system simplicity, which is much more compatible with miniaturized portable fuel-cell systems to be potentially embedded within consumer electronics devices. An early product developed for a handheld device, where a methanol cartridge and an air-breathing DMFC array were embedded by MTI Microfuel Cells to significantly enhance the device energy content, is shown in the photograph in Fig. 58. The device seen is an RFID tag reader (RFID) tag reader made by Intermec, inc., with an RF antenna located along the back, hidden surface, and an industrial version PDA resting on the handle. The DMFC cell array, located behind the vertical grill and air filter seen in the photograph breathes air through the air filter and is fed by methanol from a methanol cartridge shown in the photograph half inserted. The addition of this methanol cartridge and flat form DMFC array provided energy to the load of 35 Wh in continuous discharge to be compared with the 8 Wh provided by the rechargeable, cylindrical Li battery (located in the handle) ordinarily employed. A concept prototype of a further miniaturized DMFC and methanol cartridge considered for future integration into multifunctional handsets is shown in Fig. 6. The point to notice in Fig. 6 is the small size of the molded, DMFC array required to generate subwatt



(a)



(b)

Fig. 57 (a) Generic DMFC system utilizing neat methanol with flow and flow control subsystems devoted primarily to cathode water recovery and backpumping to an anode recirculation loop. (b) Highly simplified DMFC system developed at MTI Microfuel Cells, Albany, NY, based on direct feed of 100% methanol to the cell anode at controlled rate and elimination of any need of external water pumping.

MTI Micro's MobionTM DMFC and methanol cartridge integrated into intermec Inc.'s RFID reader



DMFC system characteristics:

Cartridge: 100% MeOH

Power out: 1 W

Energy to load: in continuous discharge: 35 Wh*

(*Present battery is 8 Wh)

Fig. 58 Air-breathing DMFC cell array and methanol cartridge integrated into the handle of an RFID tag reader made by Intermec, Inc.

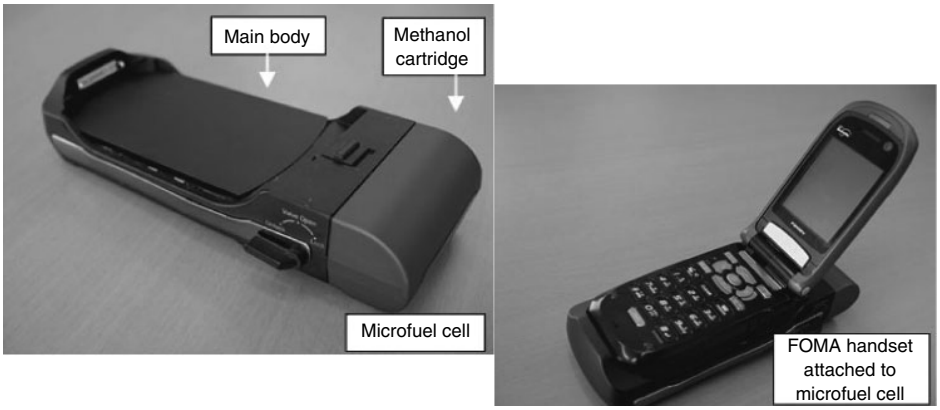


Fig. 59 A 50-W DMFC power source commercialized by Smart Fuel Cells (Germany). Volume of the 50-W (4 A; 11–14 V) stack shown is 1.3 L and it weighs 8 kg. The system efficiency is reported as 700 Wh L^{-1} (0.7 Wh cm^{-3}) of 100% methanol stored in the tank. The system noise level is reported as low as 47 dB at a distance of 1 m.

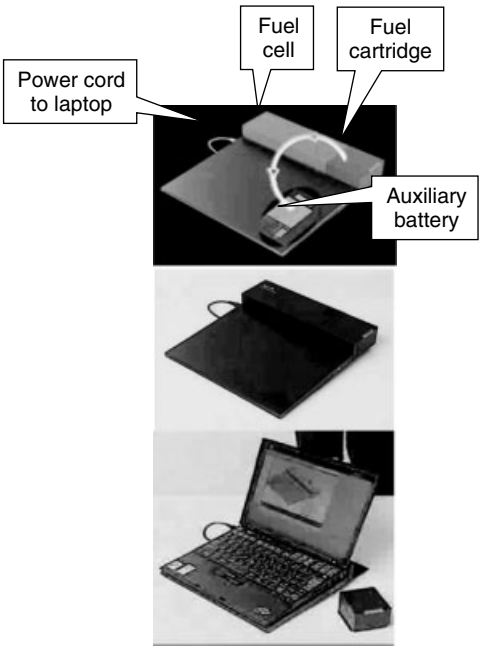
power levels. The description ''concept prototype'' stresses that the level of miniaturization depicted in this photograph is still to be achieved for a viable DMFC system of superior energy density and sufficient power output.

8.3.9.2 Demonstrations of Small Fuel-cell System Prototypes

Over the last 4–5 years, there has been a surge of interest in on microfuel cells. Many top-tier consumer electronic manufacturers have described prototype



(a)



(b)

Fig. 60 (a) Prototype microfuel cell power source developed by NTT DoCoMo (07/05) for their 3G FOMA™ handsets. Product concept shown is a charger base that cradles a mobile phone, manufactured by Fujitsu Laboratories Ltd. (b) Initial DMFC power source prototype for ThinkPads, from a joint Sanyo/IBM project (04/05). Weight with fuel reported as 2.2 kg for 8 hours. DMFC hybridized with auxiliary slim battery for power peaks up to 72 W. The volume of the prototype dock containing the fuel cell was estimated at 1600 cm^3 and the cartridge size, 160 cm^3.

fuel-cell power packs developed, with most approaches so far being based on methanol fuel – some preferring the methanol-reforming option (e.g., Motorola and Casio) and a larger number looking at DMFC-based systems (e.g., Toshiba, Samsung, and Hitachi). One manufacturer, Smart Fuel Cell (Germany), has released a portable (not man-portable) DMFC-based power source operating in the 20–50 W regime (see photo in Fig. 59). The company's first target application has been auxiliary power in recreational vehicles, as well as primary power for road-side sensors and monitors, all at a level of 20 W or higher. Apart from one case, demonstrations by Industry have focused on fuel-cell units that act basically as stand-alone sources or as chargers for the battery in an otherwise unmodified electronic device. Most industrial prototypes have been developed with laptop/notebook computers in mind (see photo in Fig. 60), for which miniaturization is a less daunting task compared to a handheld device. A common denominator for most, if not all, of demonstrated fuel-cell power packs, is that they are actually hybrid systems. The fuel cell is typically there to provide the average power over the complete period of use, whereas a relatively small battery, kept charged by the fuel cell, meets any additional peak-power requirements and, possibly, assists with system start-up at lower temperatures. Basically, the most effective way to simultaneously answer energy and power demands is to use the fuel as the energy carrier and the battery as the peak-power source. This becomes obvious from the effective energy-density advantage of fuel versus a lithium-ion battery – around $\times 5$ with current DMFC technology – and the peak maximum power-density advantage of a Li battery versus that of the DMFC – also about $\times 5$ at present. This simple observation explains why hybridization of a fuel cell (with maximum fraction of the volume reserved for fuel) with a compact battery employed to ride through periods of peak-power demand could satisfy the combined targets of more energy packaged into same volume/weight and the generation of the complete range of power demanded by the relevant load.

While at this point in time there is still no truly commercial microfuel cell power-pack product for consumer electronics devices, there is a growing consensus that low-temperature fuel cells will find their first application in providing the basis for energy-rich portable power sources either replacing or augmenting rechargeable batteries. Significantly, with the cost of incumbent battery technology pegged at several dollars per watt of power (several thousand dollars per kilowatt), the cost barrier for commercialization of fuel-cell products in this sector is much lower than in larger-scale potential applications like distributed generation and transportation.

Acknowledgments

1. I like to acknowledge more than 20 years of enjoyable collaboration with colleagues at Los Alamos National Laboratory and at MTI Microfuel Cells, doing together some of the R, D&E work described good in this chapter. Interchange of ideas and of personal expertise has always been a crucial requirement for effective advancement in this interdisciplinary area of electrochemical science and technology.
2. Padmanabhan Srinivasan of MTI Microfuel Cells performed the computer calculations of PEFC and DMFC cell characteristics for this chapter.

References

1. S. Gottesfeld, T. Zawodzinski, Polymer electrolyte fuel cells, in *Advances in Electrochemistry and Electrochemical Engineering* (Eds.: C. Tobias, H. Gerischer, D. Kolb et al.), John Wiley Publishing Company, 1997, Vol. 5.
2. S. Gottesfeld, M. S. Wilson, Polymer electrolyte fuel cells as potential power sources for portable electronic devices, in *New Trends in Electrochemical Technology: Energy Storage Systems in Electronics* (Eds.: T. Osaka, M. Datta), International Publishers Distributor (S), 2000.
3. J. Larminie, A. Dicks, *Fuel Cell Systems Explained*, Wiley, 2003.
4. (a) *Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gottesfeld, G. Halpert, A. Landgrebe), The Electrochemical Society, Pennington, NJ, 1995; (b) *Proton Conducting Membrane Fuel Cells II* (Eds.: S. Gottesfeld, T. F. Fuller), The Electrochemical Society, Pennington, NJ, 1999; (c) *Proton Conducting Membrane Fuel Cells III* (Eds.: M. Murthy, T. F. Fuller, J. W. Van Zee et al.), The Electrochemical Society, Pennington, NJ, 2002; (d) *Direct Methanol Fuel Cells, ECS Proceeding Volume No. 2001-4* (Eds.: S. R. Narayanan, S. Gottefseld, T. Zawodzinski), The Electrochemical Society, Pennington, NJ, 2001.
5. *Handbook of Fuel Cells (Four Volumes)* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), Wiley, New York, 2003.
6. See, for example, J. F. McElroy, US Patent No 6,329,090, December 2001.
7. I. D. Raistrick, in *Diaphragms, Separators, and Ion Exchange Membranes* (Eds.: J. W. Van Zee, R. E. White, K. Kinoshita et al.), The Electrochemical Society Softbound Proceedings Series, Pennington, 1986, PV 86-13, p. 172.
8. E. A. Ticianelli, C. R. Derouin, A. Redondo et al., *J. Electrochem. Soc.* **1988**, *135*, 2209.
9. E. J. Taylor, E. B. Anderson, N. R. K. Vilambi, *J. Electrochem. Soc.* **1992**, *139*, L45.
10. M. S. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1992**, *139*, L28–L30.
11. M. S. Wilson, J. A. Valerio, S. Gotetsfeld, *Electrochim. Acta* **1995**, *40*, 355.
12. M. S. Wilson, US Pat Nos. 5,211,984 and 5,234,777, 1993.
13. T. E. Springer, M. S. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1993**, *140*, 3513.
14. G. E. Mopsted, J. M. Moore, in *Handbook of Fuel Cells (Four Volumes)* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), Wiley, 2003, Vol. 3, (part 1), p. 286.
15. S. C. Yeo, A. Eisenberg, *J. Appl. Polym. Sci.* **1977**, *21*, 875.
16. R. S. Yeo, H. L. Yeager, *Mod. Aspects Electrochem.* **1985**, *16*, 437.
17. K. A. Mauritz, C. E. Rogers, *Macromolecules* **1985**, *18*, 423.
18. (a) M. Doyle, G. Rajendran, in *Handbook of Fuel Cells (Four Volumes)* (Eds.: W. Vielstich, A. Lamm, H. A. Gasteiger), Wiley, 2003, Vol. 3, p. 361; (b) see also R. S. McLean, M. Doyle, B. B. Sauer, *Macromolecules 33*, 6541 (**2000**).
19. A. Z. Weber and J. Newman, *J. Electrochem. Soc.* **2003**, *150*, A1008 and **2004**, *151*, A236.
20. La Conti in Ref.5, vol 3 (part 1), p. 647.
21. L. Xiao, H. Zhang et al., *Fuel Cells*, **2005**, *5*, 287.
22. W. Grot, G. E. Munn, P. N. Walmsley, Abstracts of the The ECS Meeting in, Houston, 1972.
23. T. A. Zawodzinski Jr, C. Derouin, S. Radzinski et al., *J. Electrochem. Soc.* **1993**, *140*, 1041.
24. J.-P. Randin, *J. Electrochem. Soc.* **1982**, *129*, 1215.
25. P. C. Rieke, N. E. Vanderborgh, *J. Membr. Sci.* **1987**, *32*, 313.
26. (a) S. J. Paddison, T. A. Zawodzinski Jr, *Solid State Ionics* **1998**, *115*, 333; (b) S. J. Paddison, L. R. Pratt, T. A. Zawodzinski Jr, *J. New Mater. Electrochem. Sys.* **1999** *2*, 183; (c) J. Paddison, L. R. Pratt, T. A. Zawodzinski Jr, *Proton Conducting Membrane Fuel Cells II* (Eds.: S. Gottesfeld, T. F. Fuller), The Electrochemical Society Proceedings Series, Pennington, 1999, p. 99 (1999).
27. K. D. Kreuer, T. Dippel, J. Maier, in *Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gotetsfeld, G. Halpert, A. Landgrebe), The Electrochemical Society, October, 1995, p. 241.
28. T. A. Zawodzinski, M. Neeman, L. Sillerud et al., *J. Phys. Chem.* **1991**, *95*, 6040.
29. T. A. Zawodzinski Jr, T. Springer, F. Uribe et al., *Solid State Ionics* **1993**, *60*, 199.
30. S. Motupaly, A. J. Baker and John W. Weidner, *J. Electrochem. Soc.* **2000**, *147*, 3171.

31. M. Eikerling, A. A. Kornyshev, Y. Kharkats et al., *J. Electrochem. Soc.* **1998**, *145*, 2684.
32. K.-D. Kreuer, T. Dippel, W. Meyer et al., *Mat. Res. Soc. Symp. Proc.* **1993**, *293*, 273.
33. X. Ren, W. Henderson, S. Gottesfeld, *J. Electrochem. Soc.* **1997**, *144*.
34. T. Fuller, J. Newman, *J. Electrochem. Soc.* **1992**, *139*, 1332.
35. T. A. Zawodzinski, J. Davey, J. Valerio et al., *Electrochim. Acta* **1995**, *40*, 297.
36. T. E. Springer, T. A. Zawodzinski, S. Gottesfeld, *J. Electrochem. Soc.* **1991**, *138*, 2334.
37. Cover page of the Oct/Nov 2005 issue of *Fuel Cell Review*, Institute of Physics Publishing.
38. C.-Y. Wang, *Chem. Rev.* **2004**, *104*, 4727.
39. D. Bernardi, M. Verbrugge, *J. Electrochem. Soc.* **1992**, *139*, 2477.
40. G. J. M. Janssen, *J. Electrochem. Soc.* **2001**, *148*, A1313.
41. (a) S. Mukerjee, S. Srinivasan, *J. Electroanal. Chem.* **1993**, *357*, 201; (b) S. Mukerjee, S. Srinivasan, M. P. Soriaga et al., *J. Phys. Chem.* **1995**, *99*, 4577.
42. T. E. Springer, T. Rockward, T. A. Zawodzinski et al., *J. Electrochem. Soc.* **2001**, *148*(1), A11.
43. US Patent No. 3,297,484 (Niedrach, 1967).
44. US Patent No. 4,272,353, June 1981 (R.J. Lawrence).
45. H. Takenaka, E. Torikai, *Kokai Tokyo Koho (Japan Patent)* **1980**, *55*, 38934.
46. P. S. Fedkiw, W.-H. Her, *J. Electrochem. Soc.* **1989**, *136*, 899.
47. P. Aldbert, F. Novel-Cattin, M. Pineri et al., *Solid State Ionics* **1989**, *35*, 3.
48. S. Srinivasan, E. A. Ticianelli, C. R. Derouin et al., *J. Power Sources* **1988**, *22*, 359.
49. E. A. Ticianelli, C. R. Derouin, S. Srinivasan, *J. Electroanal. Chem.* **1988**, *251*, 275.
50. S. Mukerjee, S. Srinivasan, A. J. Appleby, *Electrochim. Acta* **1993**, *38*, 1661.
51. M. S. Wilson, F. Garzon, K. E. Sickafus et al., *J. Electrochem. Soc.* **1993**, *140*, 2872.
52. S. C. Thomas, X. Ren, S. Gottesfeld, *J. Electrochem. Soc.* **1999**, *146*, 4354.
53. (a) M. K. Debe, in Ref. 5 Vol. 3, Chapter 45; (b) Proceedings of the Fuel Cell Workshop in Kofu, Japan, 2005.
54. J. F. McElroy, *Abstracts of the Fuel Cell Seminar*, Courtesy Associates, Phoenix, 1990, p. 282.
55. D. P. Wilkinson, J. St-Piere, in Ref. 5, Vol. 3, p. 618.
56. L. M. Roen, C. H. Paik, T. D. Jarvi, *Electrochem. Solid-State Lett.* **2004**, *7*, A19, C Corr.
57. R. M. Darling, J. P. Meyers, *J. Electrochem. Soc.* **2006**, *150*, A1523.
58. C. H. Paik, T. D. Jarvi, W. E. O'Grady, *Electrochem. Solid-State Lett.* **2004**, *7*, A82, Surface Oxidation.
59. F. Uribe, T. Zawodzinski, S. Gottesfeld, *J. Electrochem. Soc.* **2002**, *149*, A293.
60. A. B. Laconti, M. Hamdan, R. C. McDonald, in Ref. 5, Vol. 3, Chapter 49.
61. D. A. Matsen, A. D. Bosco, in Ref. 5, Vol. 4, Chapter 53.
62. D. Wheeler, J. Yi, R. Fredley et al., *J. New Mater. Electrochem. Syst.* **2001**, *4*, 233.
63. F. Barbir, in Ref. 5, p. 685, Vol. 4.
64. S. G. Chalk, J. M. Miller, F. W. Wagner, *J. Power Sources* **2000**, *86*, 45.
65. I. D. Raistrick, Internal Report, Los Alamos National Laboratory, 1985.
66. T. Springer, T. Zawodzinski, S. Gottesfeld, in *Electrode Materials and Processes for Energy Conversion and storage* (Eds.: J. Mcbreen, S. Mukerjee, S. Srinivasan), The Electrochemical Society Proceedings Series, Pennington, NJ, 1997, pp. 15–24, PV 97-13.
67. R. J. Bellows, E. Marcuchi-Soos, in *Proton Conducting Membrane Fuel Cells II* (Eds.: S. Gottesfeld, T. F. Fuller), The Electrochemical Society, Pennington, NJ, 1999, p. 218.
68. S. Gottesfeld, J. Pafford, *J. Electrochem. Soc.* **1988**, *135*, 2651–2652.
69. S. Gottesfeld, U.S. Patent No. 4,910,099 March 20, 1990.
70. P. J. Berlowitz, C. H. F. Peden, D. W. Goodman, *J. Phys. Chem.* **1988**, *92*, 5213.
71. M. S. Wilson, T. E. Springer, T. A. Zawodzinski et al., *Proceedings of the 28th Intersociety Energy Conversion Engineering Conference*, Atlanta, 1993, Vol. 1, 1203.
72. V. M. Schmidt, R. Ianniello, H.-F. Oetjen et al., in *Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gottesfeld, G. Halpert, A. Landgrebe), The Electrochemical Society, 1995, p. 1.
73. M. Iwase, S. Kawatsu, *Proton Conducting Membrane Fuel Cells I* (Eds.: S. Gottesfeld, G. Halpert, A. Landgrebe), The Electrochemical Society, Pennington, NJ, 1995, p. 12.
74. H. A. Gasteiger, N. Markovic, P. N. Ross et al., *J. Phys. Chem.* **1997**, *98*, 617.

75. N. Grgur, N. M. Markovic, P. N. Ross, in *Proton Conducting Membrane Fuel Cells II* (Eds.: S. Gottesfeld, T. F. Fuller), The Electrochemical Society, Pennington, NJ, 1999, p. 176.
76. C. M. Seymour et al., *Abstracts of the Fuel Cell Seminar*, Courtesy Associates Inc., Washington, 1992, p. 446.
77. F. Uribe, T. Zawodzinski, S. Gottesfeld, *J. Electrochem. Soc.* **2002**, *149*, A293.
78. J. Zhang, Y. Mo, M. B. Vucomiric et al., *J. Phys. Chem.* **2004**, *108*, 10955; J. Zhang, H. B. Lima, M. H. Shao et al., *J. Phys. Chem. B* **2005**, *109*, 22701–.
79. M. F. Mathias, R. Makharla, H. A. Gasteiger et al., *Interface*, The Electrochemical Society, Pennington, NJ, 2005, vol. 14(3), p. 24.
80. T. Toda, H. Igarashi, H. Uchida et al., *J. Electrochem. Soc.* **1999**, *146*, 3750.
81. H. A. Gasteiger, *Proceedings of the International Fuel Cell Workshop in Kofu*, Japan, September 2005.
82. S. Gottesfeld, I. D. Raistrick, S. Srinivasan, *J. Electrochem. Soc.* **1987**, *134*, 1455.
83. A. Parhtasarathy, C. R. Martin, S. Srinivasan, *J. Electrochem. Soc.* **1991**, *138*, 916.
84. A. Parhtasarathy, B. Dave, S. Srinivasan et al., *J. Electrochem. Soc.* **1992**, *139*, 1634.
85. A. Parhtasarathy, S. Srinivasan, A. J. Appleby et al., *J. Electrochem. Soc.* **1992**, *139*, 2856.
86. A. Parhtasarathy, S. Srinivasan, A. J. Appleby et al., *J. Electrochem. Soc.* **1992**, *139*, 2530.
87. F. A. Uribe, T. E. Springer, S. Gottesfeld, *J. Electrochem. Soc.* **1992**, *139*, 765.
88. F. Uribe, M. S. Wilson, T. E. Springer et al., *Proceedings of the Workshop on Structural Effects in Electrocatalysis and Oxygen Electrochemistry*, The Electrochemical Society Proceedings, Vol. 92-11, p. 494.
89. H. Gasteiger, W. Gu, R. Makharia et al., in Ref. 5, Vol. 3, Chapter 46.
90. T. E. Springer, T. A. Zawodzinski, M. S. Wilson et al., *J. Electrochem. Soc.* **1996**, *143*, 587.
91. H. A. Gasteiger, S. S. Kocha, B. Sompalli et al., *Appl. Catal.* **2005**, *B56*.
92. (a) A. K. Shukla, P. Stevens, A. Hamnett et al., *J. Appl. Electrochem.* **1989**, *19*, 383; (b) Data of PEFC testing at Los Alamos National Laboratory, 1986–1987.
93. G. Faubert, R. Cote, J. P. Dodelet, in *Proton Conducting Membrane Fuel Cells II* (Eds.: S. Gottesfeld, T. F. Fuller), The Electrochemical Society, Pennington, NJ, 1999, p. 30.
94. P. P. Mukherjee, C.-Y. Wang, *J. Electrochem. Soc.* **2006**, *153*(5).
95. X. G. Yang, F. Y. Zhang, A. Lubawy et al., *Electrochem. Solid-State Lett.* **2004**, *7*, A408.
96. Image courtesy of M. M. Mench, J. Benziger, Pennsylvania State University. 2006.
97. A. S. Aricò, P. Cretì, P. L. Antonucci et al., *Electrochem. Solid-State Lett.* **1998**, *1*, 66.
98. E. Peled, T. Duvdevany, A. Aharon et al., *Electrochem. Solid-State Lett.* **2000**, *3*, 12.
99. H. A. Gasteiger, N. Markovic, P. N. Ross, *J. Phys. Chem.* **1995**, *99*, 16757.
100. K. Scott, W. A. Taama, J. Cruickshank, *J. Appl. Electrochem.* **1998**, *28*, 289.
101. M. Baldauf, W. Preidel, *J. Power Sources* **1999**, *84*, 161.
102. X. Ren, T. E. Springer, S. Gottesfeld, *J. Electrochem. Soc.* **2000**, *147*(1), 92.
103. H. N. Dinh, X. Ren, F. H. Garzon et al., *J. Electroanal. Chem.* **2000**, *491*, 222–233.
104. X. Ren, M. W. Wilson, S. Gottesfeld, *J. Electrochem. Soc.* **1996**, *143*, L12.
105. M. Hickner et al., *Chem. Rev.* **2004**, *204*, 4587.
106. K. A. Mauritz, I. D. Stephanitis, S. V. Davis, *J. Appl. Polym. Sci.* **1995**, *62*, 181; also H. Uchida, Y. Hizuno and M. Watanabe, *J. Electrochem. Soc.* **2002**, *149*, A682.
107. K. D. Kreuer, in Ref.5, Vol. 3, Chapter 33.
108. DMFC hydrocarbon membrane made by Polyfuel, Mountain View.
109. X. Ren, S. Gottesfeld, US Patent No. 6,296,964.
110. D. Chu, S. Gilman, *J. Electrochem. Soc.* **1994**, *141*, 1770.
111. T. A. Zawodzinski Jr, C. Lopez, R. Jestel et al., *J. Electrochem. Soc.* **1993**, *140*, 1981.
112. G. K. S. Prakash, M. Smart, A. R. Atti et al., *The electrochemical Society Meeting*, Honolulu, 1999, Abstract 406.
113. M. Watanabe, M. H. Uchida, Y. Seki et al., *J. Electrochem. Soc.* **1996**, *143*, 3847.

9
The Electrochemistry of Nuclear Reactor Coolant Circuits

Digby D. Macdonald and Mirna Urquidi-Macdonald
Pennsylvania State University, University Park, PA, USA

9.1 Introduction **665**

9.2 Water Radiolysis **668**
9.2.1 Radiolytic Yield 669
9.2.2 Chemical Reactions 671
9.2.3 Convection 672
9.2.4 General Solution of the Equations 672

9.3 Mixed Potential Model and the Electrochemical Corrosion Potential . **673**

9.4 Stress Corrosion Cracking **678**

9.5 Electrochemistry of Boiling Water Reactors **691**

9.6 Electrochemistry of Pressurized Water Reactors **711**

9.7 Summary **719**
References 720

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

9.1 Introduction

Nuclear power reactors now account for about 16% of the world's electricity production from 440 plants with a total installed capacity of 364 000 MWe (www.uic.com.au/nip07.htm). More than 12 000 reactor years of operating experience have been accumulated. The fraction of the electrical energy generating capacity that is due to nuclear power plants varies widely, ranging from about 90% in some European nations (e.g., France and Belgium) to none in other, ''First World'' nations, such as Australia and New Zealand. While the United States of America operates the greatest number of reactors (103), compared to any nation, nuclear power accounts for only a modest 20% of its total electrical power production. However, no new reactors have been ordered in the United States for more than 30 years, although vigorous nuclear power programs have been continued elsewhere. During that time, the availability of power reactors steadily increased from less than 60% in the early 1970s to about 90% today. This is equivalent to building about one new plant per year. The incidents at Three Mile Island in the United States and at Chernobyl in the Ukraine, notwithstanding, nuclear power is now recognized as an efficient, pollution-free source of electrical energy that could supply the world's needs for many centuries to millennia, depending upon the technologies adopted. This resurgence in nuclear power is fueled by the need to reduce the emission of carbon dioxide from the burning of delayed carbon cycle (fossil) fuels (as opposed to fuels derived from the prompt carbon cycle, such as biofuels), in order to mitigate global warming, by the need to conserve the dwindling supplies of fossil fuels, and because of the political ramifications of being dependent upon foreign sources for energy, which is the ''life blood'' of nations.

The progressive development of nuclear power reactors is summarized in Fig. 1, starting with the small, prototype (Generation I) reactors to the Generation II reactors that are currently in service. In the immediate future, Generation III, advanced light water reactors (ALWRs) and advanced CANadian Deuterium oxide natural Uranium (CANDU) reactors will gradually replace the venerable Generation II plants over the time frame of 2000 to about 2030 (the first units are now operating in Japan). Beyond 2030, Generation IV reactors are expected to enter service. Generation IV reactors will have higher thermodynamic efficiencies, by virtue of their higher operating temperatures, and will employ proliferation-resistant fuels.

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

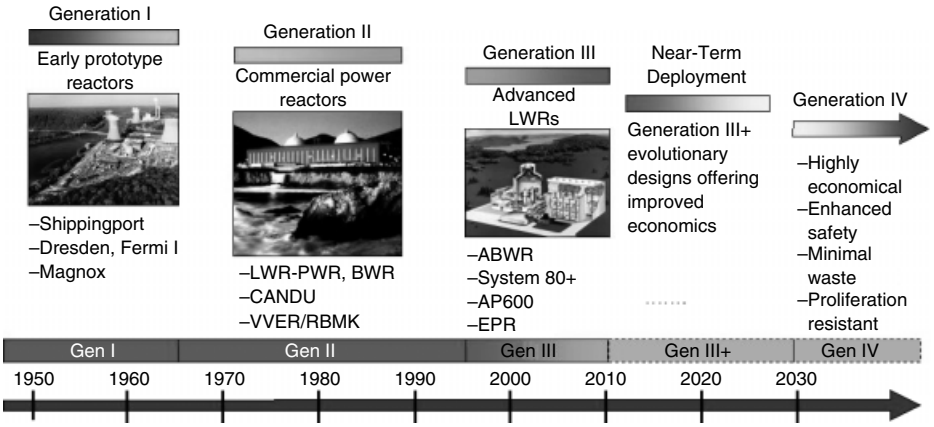


Fig. 1 Progressive development of successive generations of nuclear power reactors. http://upload.wikimedia.org/wikipedia/ commons/4/4e/GenIVRoadmap.jpg.

These plants will be ''inherently'' safe, will produce minimal waste through ''actinide burning'', and will be highly economical (but, perhaps, not so economical that the power ''will be too cheap to meter'' (Lewis L. Strauss, Speech to the National Association of Science Writers, New York City, September 16th, 1954 (New York Times, September 17, 1954))).

A common thread in many of the reactor technologies that currently exist or that are under development is the use of water as the heat transport medium (the ''coolant''). In many respects, water is an ideal coolant, because it has a high heat capacity, can be obtained in a high purity, is inexpensive, has a wide liquid range (0–374.15 °C), is easily handled, and had been used since the dawn of steam power. Thus, in their most fundamental form, water-cooled nuclear reactors (WCNRs) comprise a nuclear boiler, a heat transport system (piping, channels, steam generators, etc.), a set of turbines (high pressure, intermediate pressure, and low pressure, depending upon plant type and design), a steam condenser, and a feedwater circuit that returns the water to the core. Additional circuits include those for decay heat removal and reactor water cleanup. Thus, the ''coolant'' (water or an aqueous solution) transports the thermal energy that is produced by nuclear fission in the fuel rods located in the reactor core to either a steam generator, which generates steam in a secondary circuit, as in a pressurized water reactor (PWR, Fig. 2), or in the form of steam directly to the turbines, as in a boiling water reactor (BWR, Fig. 3). In the PWR case, the primary coolant is pressurized to prevent boiling, so that the primary coolant can be regarded as being a single-phase liquid system, whereas in the BWR case the coolant in the fuel channels is two phase (steam + water) with the water and steam being separated at the top of the reactor vessel. For thermohydraulic reasons, the water that remains in the vessel is combined with feedwater and is then recirculated through a

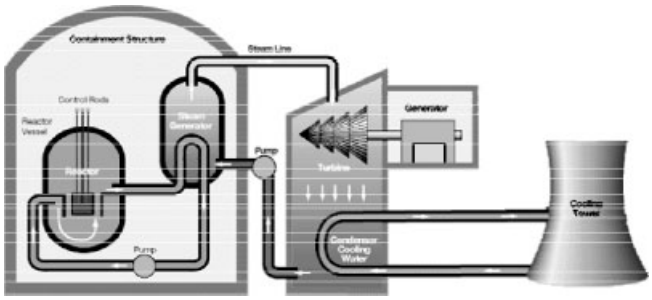


Fig. 2 Schematic of a pressurized water reactor (PWR).

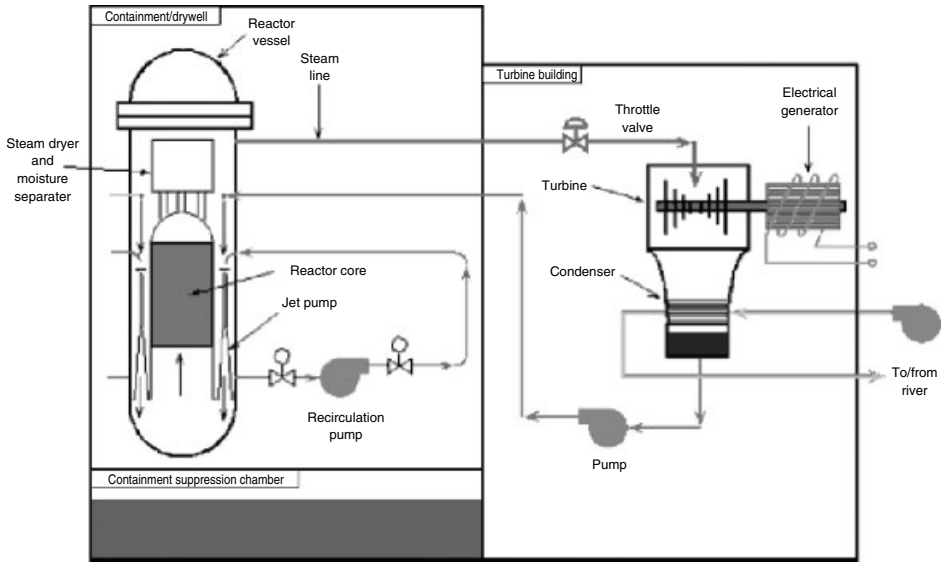


Fig. 3 Schematic of a boiling water reactor (BWR).

recirculating piping system back in to the core, so that the liquid phase can be regarded as representing a ''primary circuit'' with direct heat exchange with steam in a secondary circuit. Because boiling occurs in the core of a BWR, the coolant is nominally pure water so as to prevent the formation of concentrated solutions and deposits in boiling crevices and at heat transfer surfaces.

The temperature in the cores of modern nuclear power reactors typically ranges from 260 to about 320 °C, depending upon the reactor type and/or design (although it may be as high as 340 °C in the fuel channels if γ heating is taken into account, in the case of a PWR). Since this chapter concentrates on light water reactors (LWRs), the relevant temperatures are 288 °C for a BWR and 320 °C for a PWR. Additionally, in a reactor core, water is exposed to intense fields of ionizing radiation (α particles (He^{2+}), β particles (electrons), and γ photons), such that

water is dissociated into a myriad of species, some of which form via bond scission within the water molecule within durations that are as short as perhaps fractions of a femtosecond ($\ll 10^{-12}$ seconds), while others form via reactions involving the primary products (see below). Most of these species are electroactive (i.e., they may participate in charge-transfer reactions), and some are very strong oxidizing agents (e.g., OH), while others are strong reducing agents (e.g., H). In any event, the simultaneous presence of these species at highly nonequilibrium concentrations renders the coolant system a fertile environment for electrochemically mediated processes, such as general corrosion and various form of localized corrosion. The latter include pitting corrosion (PC), stress corrosion cracking (SCC), and corrosion fatigue (CF), phenomena that have plagued reactor cooling systems since the inception of nuclear power.

This chapter is concerned primarily with the electrochemistry of BWR and PWR primary cooling circuits and the impact that electrochemistry has on the accumulation of damage due to SCC. The available space does not permit discussion of other worthy subjects, such as PC and CF, nor does it allow the author to explore the impact of electrochemistry on the coolants in other reactor types. In keeping with the encyclopedic nature of the article, the general principles are presented and the reader is referred to the published literature for detailed information on specific topics.

9.2 Water Radiolysis

The radiolysis of water in WCNR heat transport circuits (HTCs) has long been recognized [1, 2] as a potential source of corrosive, oxidizing species, such as O_2, H_2O_2, OH, and so on, and for this reason hydrogen is added to PWR primary coolant circuits, and, in more recent years, to BWR coolant circuits, to suppress their radiolytic generation of oxidizing species and/or to displace the corrosion potentials of structural components in the coolant circuits in the negative direction. This has been accomplished in the case of PWR primary coolant by adding large amounts of hydrogen to the coolant (20–70 cc (STP)/kg, 8.93×10^{-4}–3.13×10^{-3} m, respectively), seemingly with the philosophy that ''if a little is good, a lot more must be better.'' It is argued later in this review that a lot is definitely ''not better'' with regard to the integrity of Alloy 600 that is used extensively in PWR steam generators as tube material. In the case of BWRs, the intergranular stress corrosion cracking (IGSCC) of sensitized Type 304 SS (SS = stainless steel) primary coolant components (particularly the recirculation piping system, but more recently in-vessel components) has been linked to excessively high corrosion potentials, resulting from the generation of oxidizing species (O_2, H_2O_2 due to the radiolysis of water), and to a sensitized microstructure, due to either thermal sensitization or neutron irradiation above a certain fluence (accumulated dose $> 10^{21}$ n cm^{-2}, energy > 1 MeV). The fact that the electrochemical corrosion potential (ECP) is the prime parameter in determining the accumulation of corrosion damage has been slowly accepted in the power reactor community, possibly because electrochemistry is seldom, if ever, included in mechanical and nuclear engineering curricula at universities. The impact that the ECP has on crack growth rate (CGR) is roughly exponential, if the ECP is above a critical

potential for IGSCC (E_{IGSCC}), although the rate also depends upon the stress intensity, the extent of cold working of the substrate, the conductivity of the environment, and the flow velocity of the coolant across the surface.

As noted above, the ECP is the key parameter in describing the susceptibility of reactor coolant components to corrosion damage. As experience has shown, the direct measurement of ECP in reactor coolant circuits has proved to be very difficult, notwithstanding the monumental efforts of Indig et al. at the General Electric Company-see [3], for example. The major challenge in in-reactor ECP measurements has been to devise a reference electrode that can withstand the harsh environmental conditions that exist within a reactor pressure vessel (RPV).

Electrochemists will recognize that the ECP is a mixed potential, the value of which is determined by the balance of the oxidizing and reducing species in the environment and the kinetics of dissolution (corrosion) of the substrate. In order to calculate the ECP, it is important, in principle, that the concentrations of all of the radiolytic species be determined, since all of these species are electroactive. However, theory shows that the contribution that any given species makes to the ECP is determined primarily by its concentration, so that only the most prevalent electroactive species in the system determine the ECP. This is a fortunate finding, because the various radiolysis models that are available for calculating the species concentrations do not determine the concentrations of the minor species accurately nor are there electrochemical kinetic data available for these species.

In order to calculate the species concentrations, the combined effects of the radiolytic yield of each species due to the absorption of ionizing radiation by water, and the changes in concentration due to chemical reactions and fluid convection, must be taken into account. This problem reduces to one of solving as many stiff, coupled, simultaneous first-order differential equations as there are species in the system. One of the major problems in describing the radiolysis of water is to devise efficient algorithms for performing this task.

9.2.1 Radiolytic Yield

The rate at which any primary radiolytic species is produced is given by

$$R_i^y = \left(\frac{G_i^\gamma \Gamma^\gamma}{100 N_A} + \frac{G_i^n \Gamma^n}{100 N_A} + \frac{G_i^\alpha \Gamma^\alpha}{100 N_A} \right) \tilde{F} \rho \tag{1}$$

where the rate of production of species "i", R_i^y, has units of mol cm^{-3} s^{-1}, G^n, G^γ, and G^α are the radiolytic yields for neutrons, γ photons, and α particles, respectively, in number of particles per 100 eV of energy absorbed, N_A is Avogadro's number, $\tilde{F}$ equals 6.25×10^{13} (the conversion factor from, R s^{-1} (rad s^{-1}) to eV gr^{-1} s^{-1}), and ρ is the water density in g cm^{-3}. Γ^γ, Γ^n, and Γ^α are the γ photon, neutron, and α-particle energy dose rates, respectively, in units of R s^{-1}.

Values for the radiolytic yields for various species considered in the radiolysis model for PWR coolants were taken from Christensen [4] and are listed in Table 1. A review of the literature reveals a wide variance in the G values, even from the same author. As noted by Elliot [5], the current G values should be regarded as being little more than rough estimates.

The third term in Eq. (1) accounts for the radiolysis of water by α particles

Tab. 1 Radiolytic yields (*G* values) for primary radiolytic species

Number	*Gamma*[4] *285 °C*	*Neutron*[4] *285 °C*	*Alpha*[5] *270–300 °C*	*Species*
1	1.15	0.93	0.13	e^-
2	1.08	0.50	0.12	H
3	3.97	1.09	0.45	OH
4	1.25	0.99	1.55	H_2O_2
5	0.00	0.04	0.00	HO_2
6	0.00	0.00	0.00	HO_2^-
7	0.00	0.00	0.00	O_2
8	0.00	0.00	0.00	O_2^-
9	0.00	0.00	0.00	O_2^-
10	0.00	0.00	0.00	O^-
11	0.00	0.00	0.00	O
12	0.62	0.88	1.65	H_2
13	0.00	0.00	0.00	OH^-
14	4.15	0.93	0.13	H^+

($^4He_2$ nucleus), which are produced by the $^{10}B_5(^1n_0, ^4He_2)^7Li_3$ reaction in PWR primary coolant environments, but this nuclear reaction does not occur in BWR environments, because of the absence of boron (specifically $^{10}B_5$). The importance of α-particle radiolysis in PWR cores has been recently assessed the importance of this process in PWR cores and it was concluded that the contributions from α radiolysis to the concentrations of the radiolytic species are small, when compared with those from neutrons and γ photons at the prevailing dose rates [6]. Nevertheless, in a PWR primary coolant circuit, there are regions where α-particle radiolysis contributes significantly to the formation of the radiolytic species (>10%), and hence the third term in Eq. (2) is necessary. In the case of BWR simulations, the radiolysis models [6–30] have tended to be based on the original Burns and Moore model [8], with modification being made in light of the later studies by Elliot [31]. The models used by the author and his colleagues for modeling BWR primary coolant circuits contains 34 reactions, as discussed further below.

A word of caution is needed at this point: the radiolytic yields employed in radiolysis modeling should be the *primary* yields corresponding to events that occur within the time of scission of a bond within a water molecule. These values are impossible to measure, because no techniques are available for sampling the concentrations of primary radiolysis products in the sub-femtosecond time frame. Thus, the radiolytic yields that are used to devise water radiolysis models are strictly not the values that are appropriate or needed. However, as with many physicochemical models, values for various parameters are often selected so that the original observations are reproduced. A major problem in radiolysis modeling is that there are many more parameters than there are experimental observations, so the values assigned to the models are not unique and they should not be transferred from model to model. Nevertheless, the models have proved to be quite robust, in the sense that reasonable

values for the calculated concentrations of various radiolysis products are obtained. (However, it should be noted that the concentrations of minor radiolysis products are seldom measured and compared with those that are predicted).

9.2.2 Chemical Reactions

The chemical reactions occurring in the primary HTCs of LWRs essentially determine the species concentrations in each part of the circuit, particularly in regions of low dose rate (i.e., in out-of-core regions). The reaction set used previously by Urquidi-Macdonald, Pitt and Macdonald [32] for modeling PWR primary coolant circuits has 48 reactions describing the interactions between the 14 species listed in Table 1.

This reaction set is partly based on a published compilation [8], but has been modified to include hydrogen peroxide decomposition, and the inclusion of additional species and reactions as specified by Christensen [4]. The impact of modifying the original radiolysis code, which has been used extensively in our BWR modeling work and has been extensively evaluated against BWR plant data [6, 9, 11, 17–26], is largely unknown, because no independent check on specie concentrations is currently available *for a PWR primary circuit*. This is due to the fact that the concentration of oxygen, which is monitored on a routine basis in a BWR, is so low in a PWR primary circuit that it is not measured. Furthermore, the hydrogen that is present in a PWR primary circuit is primarily the result of hydrogen additions and not radiolysis, so $[H_2]$ measured in the circuit is not a good indication of the state of radiolysis. Finally, the ECP, which is now measured on a routine basis in BWRs, is not monitored in any PWR primary circuit, to the authors' knowledge, even on an experimental basis, although some measurements have been made in VVER 440 reactors [33].

Other radiolysis mechanisms, particularly those by Christensen [4] and Elliot [31], were examined during the development of models to describe radiolysis in reactor coolant circuits, and the codes that have been developed by the author have been written to facilitate, to the greatest extent possible, the inclusion of new mechanisms. As expected, insertion of different reaction sets requires considerable reprogramming of the code.

The rate of change of each species at a given location is given by reaction rate theory as

$$R_i^c = \sum_{s=1}^{N}\sum_{m=1}^{N} k_{sm} C_s C_m - C_i \sum_{s=1}^{N} k_{si} C_s \tag{2}$$

where k_{sm} is the rate constant for the reaction between species s and m, k_{si} is the rate constant for the reaction between species s and i, and C_i, C_m, and C_s are the concentrations of species i, m, and s, respectively. N is the number of reactions in the model (i.e., $N = 48$). Explicit expressions for the gain and the loss of each species are summarized elsewhere [6].

The rate constant, k_j is a function of coolant temperature. Since the temperature throughout the heat transport circuit is not constant, the actual rate constant for each chemical reaction must be calculated for each specific position using Arrhenius' law:

$$k = k_o \exp\left[\frac{E_a}{R}\left(\frac{1}{T_o} - \frac{1}{T}\right)\right] \tag{3}$$

where k_o is the rate constant at temperature T_o, E_a is the activation energy [6], R

is the universal gas constant, and T is the temperature in Kelvin. The rate constant for hydrogen peroxide decomposition (Reaction 30) was calculated separately using an experimentally derived relationship [9]:

$$k_{30} = 1.9973e - 6\mathrm{e}^{-(14800/RT)} \quad (4)$$

Table 1 lists the species considered and the number assigned to them. Note that $[H^+]$ and $[OH^-]$ are calculated from the pH and the speciation that is afforded by the chemistry of the coolant [6].

9.2.3 Convection

As in all of the radiolysis models of reactor coolants, convection is considered to be the only mode of transport (i.e., diffusion and electromigration are neglected). In both PWR and BWR *primary* coolant circuits, single-phase (water) flow is assumed to exist in all regions of the circuit, except in the core channels of a BWR, where boiling produces the steam that carries the energy directly to the turbines. Boiling results in the transfer of volatile species (O_2 and, in particular, H_2) to the steam, which in turn reduces their presence in the liquid phase. This transfer significantly modifies the radiolysis processes in the liquid, with the result that the coolant becomes more oxidizing in nature, because of the higher rate of loss of H_2 compared with O_2 or H_2O_2. In the case of PWRs, nucleate boiling within the power-producing channels, where the local saturation vapor pressure exceeds the prevailing hydrostatic pressure, has much less effect on the coolant chemistry than does the sustained boiling in the fuel channels of a BWR, because the volatile species are returned to the liquid phase when the bubbles collapse. In nucleate boiling, the steam bubbles that form on the fuel collapse when they detach from the surface. Accordingly, any volatile radiolysis species that transfers to the steam phase is immediately returned to the primary coolant (liquid phase), so the net effect of nucleate boiling on the bulk concentrations is expected to be small, if at all it exists. This situation contrasts with that in a BWR, where a continuous steam phase is formed that leaves the primary coolant. In this case, the irreversible transfer of volatile species (H_2, O_2) to the steam has an enormous impact on the electrochemistry of the primary circuit.

9.2.4 General Solution of the Equations

The approach used in our work to solve the set of coupled ordinary differential equations (ODEs) for modeling PWR primary coolant circuits makes use of a publicly available subroutine (DVODE), which was developed by Hindmarsh at the Lawrence Livermore National Laboratory in California (www.netlib.gov). This algorithm is designed to solve first-degree, *stiff* ODE equation sets. Our system of equations is coupled throughout via the concentrations of the 14 species considered. To solve the i-coupled differential equation (i = number or species), the DVODE subroutine needs to have the set of equations and the corresponding Jacobians described. The calculation begins by assuming that the coolant is an incompressible fluid ($\bar{\nabla} \cdot \bar{v} = 0$) and that the flow is turbulent (efficient mixing). Accordingly, the flux of each dissolved species is given by

$$\overline{N}_i = -Z_i U_i F c_i \overline{\nabla}\phi - D_i \overline{\nabla} C_i + C_i \overline{v} \quad (5)$$

$$\text{Flux} = \text{migration} + \text{diffusion} + \text{convection}$$

Because of efficient mixing and the absence of an electric field, we may ignore diffusion and migration, respectively, and hence the material balance can be written as

$$\frac{\partial C_i}{\partial t} = -\overline{\nabla} \cdot \overline{N}_i + R_i \qquad (6)$$

(accumulation = net input + production), where R_i is the rate of production of the species in the fluid due to homogeneous reactions.

The primary coolant circuit of a water-cooled reactor (including BWRs and PWRs) has several loops, including the main coolant loop, a core heat removal system, and a reactor water cleanup system. However, it is convenient, for computational purposes, to differentiate between the *main* loop, which has a high flow fraction, and the secondary loops, for which the flow fractions are small. The species concentrations and electrochemical potential (ECP) are solved for in the main loop and the values at the entrance to the secondary loop are used as the initial conditions for solving the system of equations for the secondary loops of interest. Mass balance is applied at each point where more than one section comes together.

By adopting the rates of change of species mass from the various sources discussed above, we write the total rate as

$$R_i = \left(\frac{G_i^{\gamma} \Gamma^{\gamma}}{100 N_{\mathrm{A}}} + \frac{G_i^{\mathrm{n}} \Gamma^{\mathrm{n}}}{100 N_{\mathrm{A}}} + \frac{G_i^{\alpha} \Gamma^{\alpha}}{100 N_{\mathrm{A}}} \right) \tilde{F} \rho + \left[\sum_{s=1}^{N} \sum_{m=1}^{N} k_{sm} C_s C_m - C_i \sum_{s=1}^{N} k_{si} C_s \right] \qquad (7)$$

The equations must be solved iteratively until satisfactory convergence is obtained. Note that, because the mass flow rate ($\mathrm{d}m/\mathrm{d}t$) in a single (unbranched) channel is constant at all points, the linear flow rate is given by $v = (\mathrm{d}m/\mathrm{d}t)/\rho A$, where A is the cross-sectional area of the channel.

By solving Eq. (6) numerically, we are able to calculate the concentrations of each species at any point in a reactor heat transport circuit. In the actual numerical simulation, all of the parameters, except the flow velocity, circuit geometry, and temperature can be found in the published literature [6, 9, 13, 14, 34].

Transients during operation, for example, those that result from a sudden increase in power or from a step change in feedwater oxygen level, may be modeled as a series of steady states, with the initial conditions reflecting the progressive changes in the reactor's operating parameters.

The above approach describes the method used by Urquidi-Macdonald, Pitt and Macdonald [32] in modeling PWR primary coolant circuits, but that employed by Yeh, Motta, and Macdonald [6, 13, 14], Yeh et al. [20–29], and others [11, 12, 15–17] for modeling BWR primary coolant circuits is somewhat different, but yields similar results. The reader is referred to the original literature for details.

9.3 Mixed Potential Model and the Electrochemical Corrosion Potential

After the concentration of each radiolysis species is calculated, the corrosion potential of the component can be estimated using a mixed potential model (MPM) [35]. The MPM is based on the physico-electrochemical condition that charge conservation must be fulfilled at the corroding interface. Because electrochemical reactions transfer charge across a metal/solution interface at rates measured by the partial currents, the following

equation expresses the charge conservation constraint:

$$\sum_{j=1}^{n} i_{\mathrm{R/O},j}(E) + i_{\mathrm{corr}}(E) = 0 \qquad (8)$$

where $i_{R/O,j}$ is the partial current density due to the jth redox couple in the system and i_{corr} is the metal oxidation (corrosion) current density. These partial currents depend on the potential drop across the metal/solution interface.

In the current version of the MPM, which was developed for modeling the ECP of Type 304 SS in BWR primary circuits, the steel oxidation current density, i_{corr}, was modeled as an empirical function of voltage, based on the data of Lee (see Ref. 36),

$$i_{\mathrm{corr}} = \frac{\mathrm{e}^{(E-E_{\mathrm{o}})/b_{\mathrm{f}}} - \mathrm{e}^{-(E-E_{\mathrm{o}})/b_{\mathrm{r}}}}{384.62\mathrm{e}^{4416/T} + X} \qquad (9)$$

where

$$X = \frac{\mathrm{e}^{(E-E_{\mathrm{o}})/b_{\mathrm{f}}}}{2.61 \times 10^{-3}\mathrm{e}^{-4416/T+0.523(E-E_{\mathrm{o}})^{0.5}}} \qquad (10)$$

and

$$E_{\mathrm{o}} = 0.122 - 1.5286 \times 10^{-3}T \qquad (11)$$

In these expressions, b_{f} and b_{r} are the forward and reverse Tafel constants, respectively, for the metal dissolution reaction, with values of 0.06 V being assumed for both. Actually, they are empirical constants that were assumed a priori in fitting Eq. (9) to the current/voltage data. It is important to note that Eq. (9) applies strictly to Type 304 SS in near neutral solutions [35] and hence that this expression may not be a good empirical model for stainless steels in PWR primary circuits. More recently, the point defect model (PDM) [37] has been used as the basis for calculating the corrosion current density of the metal substrate.

The PDM yields the passive current density in the form

$$I_{\mathrm{ss}} = \Gamma F\left[k_2^0 \mathrm{e}^{\alpha_2\alpha\gamma\mathrm{V}}\mathrm{e}^{-\alpha_2\alpha\varepsilon\gamma L_{\mathrm{ss}}} + k_4^0\mathrm{e}^{\alpha_4\alpha\gamma\mathrm{V}}\mathrm{e}^{\alpha\beta\gamma\mathrm{pH}} + k_7^0 \cdot C_{\mathrm{H^+}}^n\right] \qquad (12)$$

where the film thickness is given by

$$L_{\mathrm{ss}} = \frac{1-\alpha}{\varepsilon}\mathrm{V} + \frac{1}{\varepsilon}\left\{\frac{2.303\mathrm{n}}{\alpha_3\chi\gamma} - \beta\right\}\mathrm{pH} + \frac{1}{\alpha_3\chi\mathrm{K}}\ln\left(\frac{k_3^0}{k_7^0}\right) \qquad (13)$$

The parameters in these equations are as follows: χ, α, β, and n are the oxidation state of the cation in the barrier layer; the polarizability of the film/solution interface (i.e., the dependence of the potential drop across the film/solution interface on the applied voltage); the dependence of the potential drop across the film/solution interface on the pH; and the kinetic order of the film dissolution reaction with respect to hydrogen ion concentration, respectively. Note that, in deriving Eqs (12) and (13), the oxidation state of the cation in the barrier layer (χ) is set equal to the oxidation state of the same cation in the solution/outer layer. The standard rate constants, k_i^0, and α_i correspond to the reaction shown in Fig. 4, ε is the electric field strength, $\gamma = F/RT$, and $K = \varepsilon\gamma$. The three terms on the right side of Eq. (12) arise from the transmission of cation interstitials, the transmission of cation vacancies, and the transmission of oxygen vacancies (or dissolution of the film), respectively. Values for these parameters are readily obtained by optimizing the PDM on

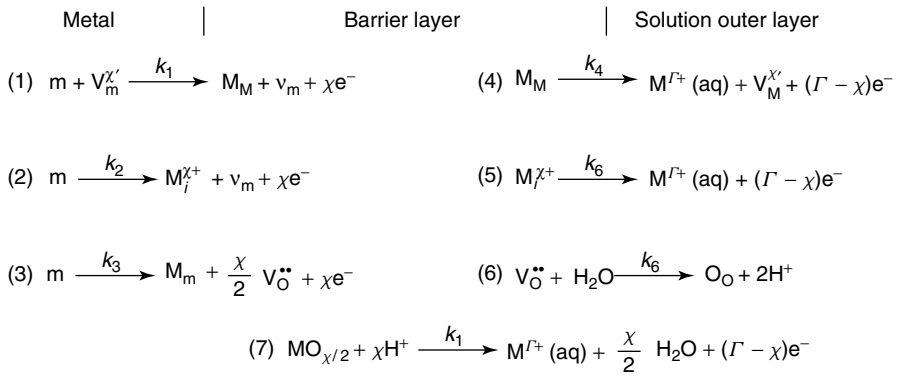


Fig. 4 Summary of the defect generation and annihilation reactions envisioned at the interfaces of the barrier oxide layer on a metal, according to the PDM [37]. $V_M^{\chi'} \equiv$ cation vacancy in the film; $\nu_m \equiv$ vacancy in the metal substrate; $M_i^{\chi+}$ cation interstitial; $V_O^{\bullet\bullet} \equiv$ oxygen (anion) vacancy; $M^{\Gamma+}(aq) \equiv$ cation in solution; M_M cation in cation site on the metal sublattice; O_O oxide ion in anion site on the oxygen sublattice; and $MO_{\chi/2} \equiv$ stoichiometric barrier layer oxide. Note that Reactions 1, 2, 4, 5, and 6 are lattice conservative processes (they do not result in the movement of the interface), whereas Reaction 3 and 7 are nonconservative.

electrochemical impedance spectroscopic (EIS) data [38].

Because electrochemical kinetic data are available only for the hydrogen electrode reaction (HER, H_2/H^+), the oxygen electrode reaction (OER, O_2/H_2O), and the hydrogen peroxide electrode reaction (HPER, H_2O_2/H_2O), only H_2, O_2, and H_2O_2 are considered as the redox species in the MPM. The redox reactions of interest are therefore written as follows [35]:

$$H^+ + e^- \longrightarrow 1/2H_2 \quad \text{or}$$
$$H_2O + e^- \longrightarrow 1/2H_2 + OH^- \quad (14)$$
$$O_2 + 4H^+ + 4e^- \longrightarrow 2H_2O \quad \text{or}$$
$$O_2 + 2H_2O + 4e^- \longrightarrow 4OH^- \quad (15)$$
$$H_2O_2 + 2H^+ + 2e^- \longrightarrow 2H_2O \quad \text{or}$$
$$H_2O_2 + 2e^- \longrightarrow 2HO^- \quad (16)$$

depending upon the pH. For pH values of less than about 4, the acid forms of the reactions prevail (left side), while at higher pH values the base forms predominate. The exact form of the reaction has an important impact on the MPM, because it determines the reaction order with respect to the concentration of hydrogen ion in the expression for the exchange current density [17].

The current density ($i_{R/O}$) for a redox couple (e.g., O_2/H_2O, H^+/H_2, H_2O_2/H_2O)

$$R \Longleftrightarrow O + ne \quad (17)$$

(where R is the reduced species and O is the oxidized species) can be expressed in terms of a generalized Butler–Volmer equation as

$$i_{R/O} = \frac{e^{(E-E^e_{R/O})/b_a} - e^{-(E-E^e_{R/O})/b_c}}{\frac{1}{i_{0,R/O}} + \frac{1}{i_{l,f}} e^{(E-E^e_{R/O})/b_a} - \frac{1}{i_{l,r}} e^{-(E-E^e_{R/O})/b_c}} \quad (18)$$

where $i_{0,R/O}$ is the exchange current density, $i_{l,f}$ and $i_{l,r}$ are the mass-transfer limited currents for the forward and reverse directions of the redox reaction, respectively, and b_a and b_c are the anodic and cathodic Tafel constants. $E^e_{R/O}$ is the equilibrium potential for this reaction, as computed from the Nernst equation:

$$E^e_{O/R} = E^0_{O/R} - \frac{2.303RT}{nF}\log\left(\frac{a_R}{a_O}\right) \quad (19)$$

where a_R and a_O are the thermodynamic activities of R and O, respectively, and $E^0_{O/R}$ is the standard potential. Limiting currents are calculated using the equation:

$$i_{l,O/R} = \pm 0.0165 nFDC^b_{O/R} Re^{0.86} Sc^{0.33}/d \quad (20)$$

where the sign depends on whether the reaction is in the forward (+) or reverse (−) direction, F is Faraday's number, D is the diffusivity of the redox species, $C^b_{O/R}$ is the bulk concentration of O or R, as appropriate, Re is the Reynolds number ($Re = Vd/\eta$), Sc is the Schmidt number ($Sc = \eta/D$), d is the channel diameter, V is the flow velocity, and η is the kinematic viscosity.

An important point that needs to be emphasized again is that the maximum contribution that any given radiolytic species can make to the ECP is roughly proportional to its concentration. Thus, in BWR simulations, the concentrations of H_2, O_2, and H_2O_2 are calculated to be orders of magnitude greater than those of any other radiolytic species and hence only these three need be considered. In the case of PWR primary HTCs, our previous modeling [26] suggests that equated electrons, H atoms, and OH radicals may be significant species in regions of very high energy dose rate (e.g., near the fuel). However, no electrochemical kinetic data exist for these reactions and hence they cannot be incorporated at present.

The accuracy of the MPM has been evaluated by comparing calculated ECP values for Type 304 SS against measured BWR plant data. While modeling of the ECP in BWR primary coolant circuits will be discussed at length later in this chapter, a brief discussion of some of the data is appropriate at this point. Thus, the most comprehensive comparison is probably that which employed data obtained during a hydrogen water chemistry (HWC) mini-test at the Leibstadt BWR in Switzerland. The reactor water chemistry modeling and the prediction of the ECP was carried out in a "double-blind" manner (to the plant $[O_2]$, $[H_2]$, or ECP data prior to submission of our calculations and the contractor did not have access to our calculations while performing the mini-test) [19]. We were, of course, provided with the flow velocity, hydrodynamic diameter, and temperature data for the test system. The calculated and measured (plant) ECP data for this case are summarized in Fig. 5. Excellent agreement is obtained in systems to which hydrogen had been added, with the measured and calculated ECP values agreeing within the combined uncertainty levels. In the normal water chemistry case ($[H_2] = 0$), the measured ECP is significantly higher than the calculated value. This is almost certainly due to the presence of hydrogen peroxide in the coolant, which was not measured by the personnel conducting the HWC mini-test. Accordingly, we were unable to input a value for $[H_2O_2]$ into the MPM. However, if we use the calculated values for $[H_2]$, $[O_2]$, and $[H_2O_2]$ obtained from RADIOCHEM [9], excellent agreement is obtained.

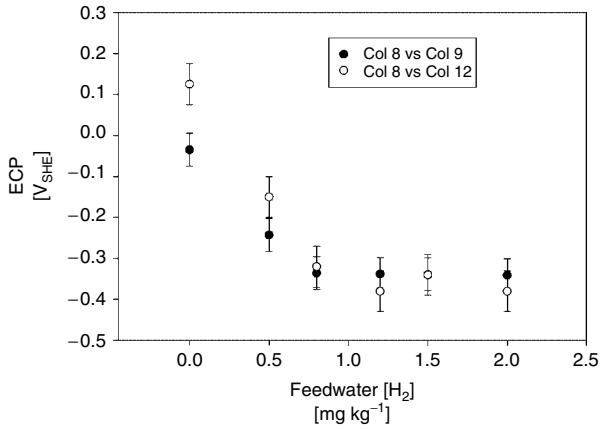


Fig. 5 Comparison of calculated and measured ECP for the Leibstadt BWR. The ''measured'' data were acquired in a test cell attached to the recirculation piping.

At this point, it is of interest to explore the impact that oxygen and hydrogen peroxide will have on the ECP under typical PWR primary circuit conditions. This is done in Fig. 6 for the system parameter values as summarized in the caption. In the case of both oxygen and hydrogen peroxide, the ECP is predicted to be displaced in the positive direction for $[O_2]$ or $[H_2O_2]$ greater than 0.001 mg kg^{-1} (1 ppb). Below this concentration, the ECP is dominated by the HER. At high oxidant concentrations (>0.1 mg kg^{-1}), the stronger oxidizing power of H_2O_2 becomes evident, in that the ECP is progressively displaced to more positive

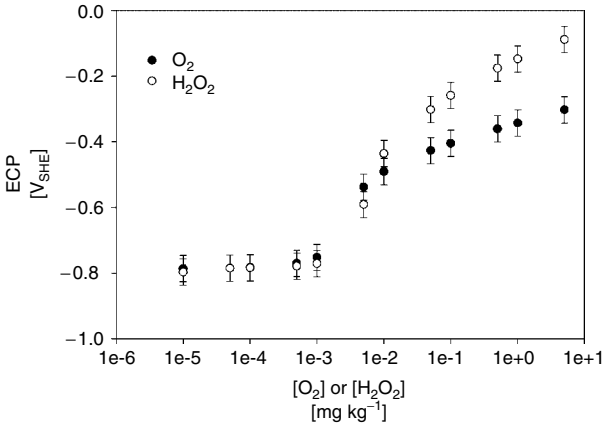


Fig. 6 Calculated ECP versus $[O_2]$ or $[H_2O_2]$ for a simulated PWR coolant system. $T = 320\,^{\circ}C$, $[H_2] = 25$ cc (STP)/kg (H_2O), $[B] = 1000$ mg kg^{-1}, $[Li] = 2$ mg kg^{-1}, pH = 7.36, flow velocity = 100 cm s^{-1}, channel hydrodynamic diameter = 100 cm.

values compared with that for oxygen as the oxidant concentration is increased. The critical oxidant concentration of 1 μg kg^{-1} (1 ppb) at which the ECP is displaced from the hydrogen electrode behavior is in good semiquantitative agreement with experimental data obtained from laboratory studies by Bertuch et al. [39].

Finally, we note that, at least for BWR coolant conditions, and possibly for PWR primary conditions, also, when the ECP is not ''clamped'' by the HER, the corrosion potential is a sensitive function of the hydrodynamic conditions in the channel. This sensitivity to flow arises because the hydrodynamic conditions control the rate of mass transport of the redox species to the metal surface. This is an extremely important issue when comparing the calculated ECP with laboratory and plant data, because the measurements are seldom carried out in a channel that is well-characterized hydrodynamically. Indeed, a well-characterized channel should be a prerequisite for performing measurements of this type.

9.4 Stress Corrosion Cracking

The primary motivation for predicting the electrochemical properties of the coolant circuits of water-cooled nuclear power reactors has been that of explaining and predicting tenacious operating problems that include SCC and CF, mass transport of corrosion products and subsequent fouling of heat transfer surfaces, activity transport due to the movement of neutron-activated radionuclides from the core to out-of-core surfaces that are not shielded, and, in the case of PWRs, the axial offset anomaly (AOA). This latter phenomenon results from the deposition of boron compounds (e.g., $LiBO_2$) in CRUD on the fuel surfaces, which absorb neutrons, and leads to uncontrolled and unplanned reductions in reactor power. The resulting γ photon fields from the transport and deposition of neutron-activated products exact a high man-REM cost on the plant operators, thereby requiring expensive shielding and/or personnel replacement. Space does not permit a discussion of all of these topics here, so the discussion will be restricted to IGSCC under BWR environmental conditions. It is fair to say that the theory for SCC is much better developed in the case of BWRs than in the case of PWRs, no doubt reflecting the more urgent nature of cracking in the former than in the latter.

IGSCC of heat-affected zones (HAZs) adjacent to welds in Type 304 SS components in the primary circuits of BWRs have plagued the industry for more than 30 years. The cracks first appeared in the 1970s in small-diameter pipes in the recirculation system, but eventually spread to the main recirculating piping, thereby raising concern of a possible loss of coolant accident (LOCA). Cracking eventually spread to in-vessel components, including the inner surface of the shroud, jet pump hold-down beams, and control rod guide tubes, for example. The cause of cracking is a combination of having a susceptible (sensitized) microstructure and a corrosive environment, coupled with a tensile stress. Thermal sensitization results from the reaction of carbon and chromium to precipitate chromium carbides (e.g., $Cr_{23}C_7$) on the grain boundaries as the temperature falls through the 800–500 °C temperature range in a region of material close to the weld fusion line. The precipitation of the carbides results in the formation of a chromium-depleted zone, within which the chromium concentration

has fallen below the 12% that is necessary to maintain the steel ''stainless.'' Thus, the grain boundaries become susceptible to stress-assisted corrosion, accounting for the intergranular morphology of the fracture. There is no question that cracking was exacerbated by the choice of Type 304 SS, which has carbon levels of the order of 0.08%, rather than of the low carbon grade, Type 304L SS (<0.02% C), but the mechanical properties of the latter were insufficient to meet the design codes. Furthermore, the susceptibility of the high carbon grade to IGSCC in high-temperature water was, perhaps, not fully appreciated. However, it is informative to note that, in the 1960s, the Germans chose the Nb-stabilized grade, Type 347 SS, for fabricating the heat transport system piping of their internal-pump BWRs, because of its greater resistance to sensitization and hence to IGSCC.

Given the seriousness of IGSCC in BWR primary coolant systems, it is not surprising that various models have been developed for describing this phenomenon, including PLEDGE by Andresen and Ford [40], the coupled environment fracture model (CEFM) of Macdonald and coworkers [41–46], the finite element model (FEM) of Gavrilov, M. Vankeerbergen, G. Nelissen and J.Deconinck [47], and the mechanical model of Peng, Kwon and Shoji [48], to name a few. It is important to note that the experimental evidence clearly shows that IGSCC in sensitized Type 304 SS under BWR environmental condition is primarily an electrochemical phenomenon and any model that does not recognize this fact cannot succeed in describing the known, strong dependencies of the CGR on the various ''electrochemical'' independent variables, including the ECP, environment conductivity, pH, and flow velocity, and for the weak dependence on stress intensity in the Stage II region of the CGR versus K_I correlation. These observations essentially rule out Peng et al.'s [48] treatment as being viable. Likewise, PLEDGE [39] is essentially empirical in nature and, while it is capable of predicting CGR within the field of calibration, empirical models are generally poor at predicting the dependent variable beyond this field [27]. Both the CEFM and FEM are ''deterministic'' models, in that they are based upon sound physical crack advance mechanisms *and* the predictions are constrained by the natural laws, in this case the conservation of mass and charge. These constraints ensure that the solution to the problem is that which is consistent with our integrated scientific experience. The principal difference between the CEFM and the FEM is that the CEFM (at least later versions [46]) provides an analytical solution to the problem of calculating the CGR, which is fast in its execution, whereas generally the FEM must employ a fine mesh in order to capture the details of the crack geometry and the exterior surfaces, and hence is much slower. In the prediction of accumulated damage over a defined corrosion evolutionary path (CEP) of the reactor, the CGR algorithm might be called thousands of times, so execution speed is a matter of prime importance. It is partly for this reason that the analytical (but approximate) version of the CEFM developed by Engelhardt, Macdonald, and Urquidi-Macdonald [46] for describing and predicting IGSCC in sensitized Type 304 SS in high-temperature aqueous solutions.

As noted above, the CEFM was developed to accurately predict the accumulation of SCC damage during the start-up and shutdown of BWRs, when the chemical and electrochemical conditions within the coolant change sharply with time. In

developing one version of the model [45], it was also necessary to introduce temperature dependence into the crack-tip strain rate, in the form of the Arrhenius equation. The resulting algorithm accurately predicts CGR in sensitized Type 304 SS over a wide temperature range (50–300 °C).

The CEFM is based upon the differential aeration hypothesis (DAH) for localized corrosion that was first enunciated by U. R. Evans at Cambridge University in the 1920s. The DAH postulates that there exists a spatial separation between the local anode, which is located in that region of the system that has the least access to the cathodic depolarizer (e.g., O_2), and the local cathode, which forms at the location that has the greatest access to the cathodic depolarizer, with the proviso that both regions must be in electronic and electrolytic communication (Fig. 7). This arrangement results in the flow of positive current in the electrolyte phase from the crack to the external surface and a corresponding net electron current flowing from the crack tip to the external surface through the metal. This electron current on the metal side of the interface annihilates an equal positive current on the solution side of the interface via a charge-transfer reaction (e.g., the reduction of oxygen).

Measured CGR data for Type 304 SS in dilute sulfuric acid solutions at various temperatures show a maximum in the CGR versus temperature at around 175 °C. Previously, this temperature-dependent CGR behavior has been modeled using the CEFM without a temperature-dependent crack-tip strain rate, with any temperature dependence being attributed to the electrochemical reactions that occur on the external surface and at the crack tip and to the transport properties of ions in the crack-internal and external environments.

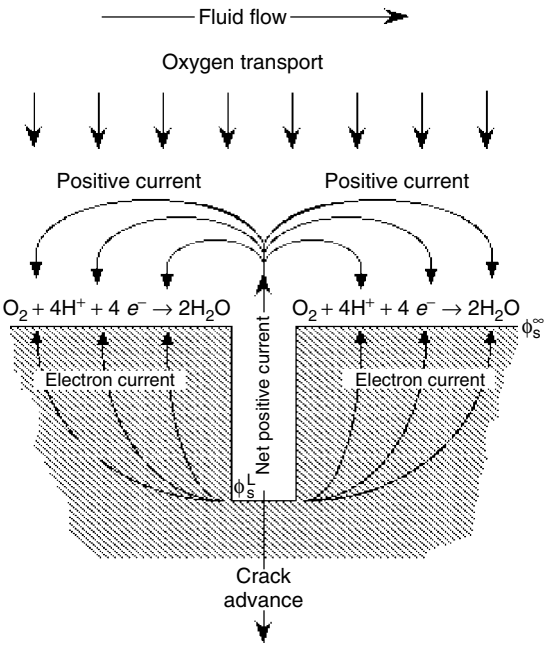


Fig. 7 Coupling of crack-internal and crack-external environments. Note that in the steady state, the crack can grow only as fast as the positive current flowing from the crack can be consumed on the external surfaces by oxygen reduction.

The version of the CEFM developed by Vankeerbergen and Macdonald [45] assumed dilute sulfuric acid solution as the reaction medium, because this is the medium for which experimental CGR versus temperature data are available from the work of Andresen [49]. Four ionic species are present in dilute sulfuric acid solution, to a significant extent: the hydrogen ion H^+, the hydroxyl anion OH^-, the bisulfate anion HSO_4^- and the sulfate anion SO_4^{2-}. The bulk concentrations of these species are readily calculated as a function of temperature by using a speciation model, thereby yielding the conductivity of the solution. Parenthetically, it is assumed that the first dissociation of sulfuric acid ($H_2SO_4 \Leftrightarrow H^+ + HSO_4^-$) is complete and that the concentrations of the species may be calculated to sufficient accuracy using the dilute solution theory. Details of the calculation can be found in Ref. 45. Furthermore, the electrode (''corrosion'') potential, E_{corr}, at a sufficiently large distance away from the crack mouth such that the crack has no impact on its value, is calculated using the MPM [35] described above. Here, for Type 304 SS in a dilute sulfuric acid solution, only oxygen reduction and metal dissolution are considered as the partial electrode reactions, although previous modeling work has also included hydrogen evolution. Calculation of E_{corr} is the first step in the CEFM algorithm.

The potential and current density in the external environment is modeled by applying charge conservation and Ohm's law to yield Laplace's equation

$$\nabla^2\varphi = 0 \quad and \quad i = -\sigma\nabla\varphi \qquad (21)$$

where ϕ is the electrostatic potential in the solution, i is the current density, and σ is the conductivity (Fig. 7). The boundary conditions are the mixed potential at an effectively infinite distance away from the crack mouth, E_{corr}, and the potential at the crack mouth, E_{mouth}. A constant, but temperature-dependent, conductivity, σ, is assumed, owing to the four ionic species in the solution (H^+, OH^-, HSO_4^-, SO_4^{2-}). By solving Laplace's equation in the external environment (given E_{corr}, E_{mouth}, and σ), the 'external' current at the crack mouth, I_{ext}, is obtained.

Inside the crack (Fig. 10), the transport of various ions, i, is modeled using the Nernst–Planck equation

$$j_i = \frac{-F}{RT} z_i D_i C_i \nabla\varphi - D_i \nabla C_i \qquad (22)$$

where j_i is the flux associated with charged species i, R is the universal gas constant, F is Faraday's constant, T is the Kelvin temperature, D_i is the diffusion coefficient, z_i is the charge, and C_i is the concentration. In previous versions of the CEFM [41–44], the diffusion-migration equations are simplified by neglecting crack wall currents and assuming a linear potential profile down the crack, as dictated by the solution to Laplace's equation in one dimension. Agreement between experimental and modeling results shows that ignoring crack wall currents is not critical (see later).

It is assumed that the slip dissolution mechanism [40] adequately describes the crack-tip process. The controlling variables are the stress intensity factor (from mechanical loading) and the crack-tip electrode potential (from electrochemical loading). The crack-tip repassivation process is important, because the kinetics of repassivation determine the fraction of the crack-tip area that remains ''bare'' over a slip-dissolution-repassivation cycle. The temperature dependence of the crack-tip process is brought into play through a temperature-dependent crack-tip strain

rate. An Arrhenius temperature dependence is assumed

$$\dot{\varepsilon}(T) = \dot{\varepsilon}(288\,^{\circ}\mathrm{C}) \exp\left\{-\frac{Q}{R}\left(\frac{1}{T} - \frac{1}{561.15}\right)\right\} \quad (23)$$

where $\dot{\varepsilon}\,(T)$ and $\dot{\varepsilon}\,(288\,^{\circ}\mathrm{C}, 561.15\mathrm{K})$ are the crack-tip strain rate at temperature T and 288 °C, respectively, and Q is the thermal activation energy for the crack-tip strain rate.

In this numerical version of the CEFM, a potential E_{mouth} is assumed at the crack mouth, the boundary between the crack-internal and the crack-external environments, and this potential is then iterated until the external current (I_{ext}, obtained via solving Laplace's equation) and the internal current (I_{int}, obtained via solving the Nernst–Planck equations) match for the prevailing E_{corr} (Fig. 8), that is, until $I_{\mathrm{int}} = I_{\mathrm{ext}}$. This is the practical implementation of the CEFM statement that the governing system equation is charge conservation,

$$i_{\mathrm{crack}} \cdot A_{\mathrm{crackmouth}} + \int_S i_{\mathrm{C}}^{\mathrm{N}} \mathrm{d}s = 0 \quad (24)$$

where i_{crack} is the net (positive) current density exiting the crack mouth, $A_{\mathrm{crackmouth}}$ is the area of the crack mouth, $i_{\mathrm{C}}^{\mathrm{N}}$ is the net (cathodic) current density due to charge-transfer reactions on the external surface, and ds is an increment in the external surface area. The subscript "S" on the integral indicates that the integration is to be performed over the entire external surface. Iteration on the crack mouth potential is necessary, rather than on the corrosion potential (ECP) as suggested elsewhere [50], because the point at which the corrosion potential is established is sufficiently remote from the crack so that no impact of the crack is evident and because the ECP is fixed for fixed external conditions.

To solve the simplified Nernst–Planck equations in the internal crack environment, it is common practice in the numerical version of the CEFM to assume an electrode potential E_{tip} at the crack tip and to iterate on this potential until electroneutrality is observed at that location (Fig. 9):

$$\sum_i z_i C_i = 0 \quad (25)$$

where z_i is the charge and C_i the concentration of species i. The choice of the crack-tip location for this iteration is somewhat arbitrary and one might argue that the iteration should be performed at all locations along the crack. However, the latter strategy would require significantly greater computation time and hence would

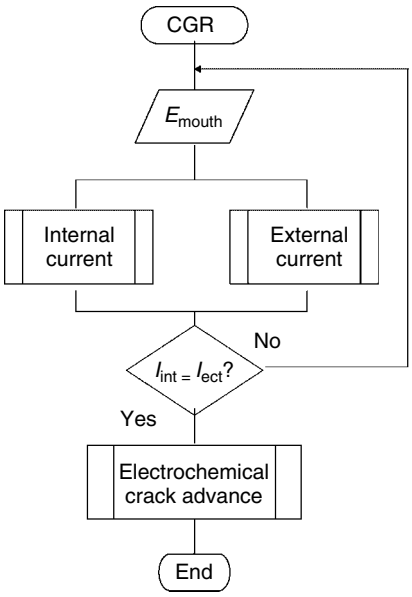


Fig. 8 Charge conservation – iteration on the mouth potential, E_{mouth}.

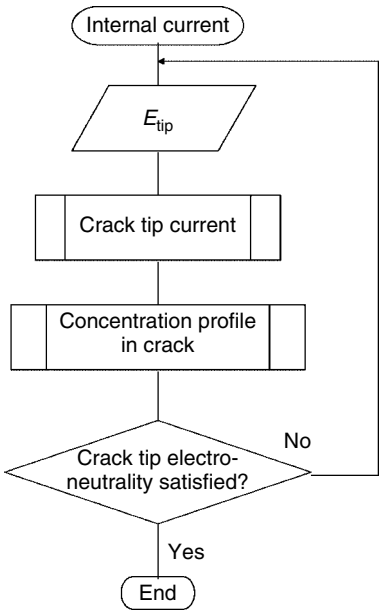


Fig. 9 Electroneutrality – Iteration on the crack-tip potential, E_{tip}.

negate the principal attribute of the CEFM: it is fast (although not nearly as fast as the approximate analytical model [46]).

It is well known that, in the fracture of sensitized stainless steels and nickel alloys in oxidizing aqueous media, the CGR becomes independent of potential at sufficiently negative ECP values [40, 43] and that the fracture morphology changes from intergranular brittle fracture to ductile failure. The ductile fracture surfaces frequently yield evidence of microvoid coalescence, with ductile tearing of the matrix between the voids. These voids appear to nucleate at intergranular precipitates, such as carbides (e.g., $Cr_{23}C_7$) on the grain boundaries, but nucleation at intragranular precipitates is also observed. Thus, in the CEFM, which was developed originally to describe IGSCC in sensitized stainless steels, it seemed appropriate to describe crack growth at the "creep" limit in terms of a cavitation model. A modified version of the cavitation model developed by Wilkinson and Vitek [51] was used to estimate CGR at sufficiently negative potentials, where environmental effects are not evident.

The treatment of creep crack growth via cavity nucleation and linkage assumes that the crack advances through a series of microfracture events at the crack front [38]. Assuming that the dimension of such an event is c, the CGR is derived (in one dimension) as

$$\frac{da}{dt} = fc \tag{26}$$

where f is the microfracture event frequency [51, 52].

Parameter values used in this model are given in the original Refs 45, together with the sources from which they were taken. For $K_I = 27.5$ Mpa.m$^{1/2}$, the CGR in Type 304 SS at 288 °C was estimated to be 1.67×10^{-10} cm s^{-1}, which was in good agreement with the observed, potential-independent CGR for ECP values more negative than ~ -0.3 V_{SHE} [40, 43]. An accurate estimate of the creep CGR is particularly important, because it defines the lower limit to the rate of crack advance and is instrumental in defining the critical potential, E_{IGSCC} (potential at which the stress corrosion CGR is observed to exceed the creep rate).

The theoretically estimated and experimentally determined CGRs agree well over a considerable temperature range after calibration at one temperature and specification of an appropriate activation energy for the crack-tip strain rate (Fig. 10). The numerical solution employed in the version of the CEFM described here yields very reasonable results for the environmentally assisted and creep fracture of sensitized

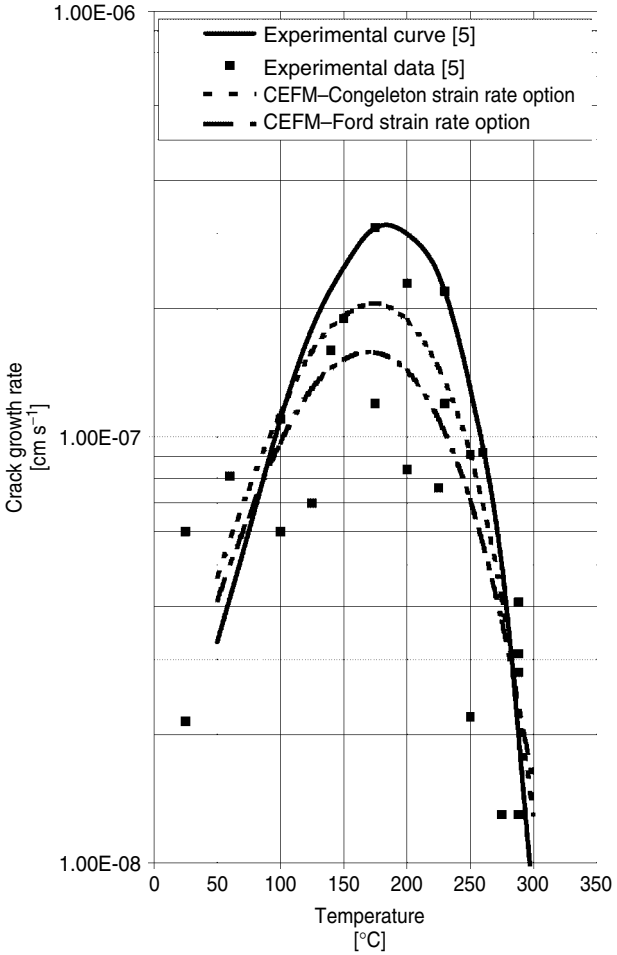


Fig. 10 The effect of temperature on crack growth rate in Type 304 stainless steel in dilute sulfuric acid solution (experimental data taken from Andresen [49] as cited in Ref. 45).

Type 304 SS, and has yet to yield a prediction of CGR that is at odds with experiment (i.e., lies outside of the experimental error range). The calibration factor appears to take care of a number of simplifying assumptions in the numerical CEFM (e.g., inert crack walls and linear potential drop down the crack) and compensates for some less quantified effects (e.g., parameters associated with the crack-tip process). Nevertheless, it is only necessary to calibrate the model with a single CGR at specified conditions and to choose an appropriate activation energy for the crack-tip strain rate (this is actually equivalent to calibrating the model with CGR data at two temperatures). As noted above, an approximate analytical solution to the environmentally assisted cracking (EAC) problem has

been published [46] and a finite element algorithm, which captures great detail of the crack propagation process, has also been described [47]. The latter approach, in principle, enables more complex geometries to be considered by moving the approach into a more efficient numerical computational framework, albeit at a considerable cost of computational time compared with the analytical solution.

One of the persistent issues that is raised, in debates on the basis of the CEFM, concerns the role of the external environment in determining the CGR. If the DAH is accepted as the basis for localized corrosion, including SCC, then the importance of the external environment is unequivocal. That the external environment is intimately involved has been demonstrated experimentally by detecting and measuring the coupling current that flows from the crack to the external surfaces where it is consumed by oxygen reduction [41]. If the cathodic processes were confined to the crack, the measured coupling current would be zero, which would be at odds with experiment [52]. Thus, the conservation of charge *requires* consideration of the external environment in any deterministic description of crack growth. Indeed, the event that alerted the author to the essential role played by the external surfaces was a failed SCC experiment at the Ohio State University, while he was Director of the Fontana Corrosion Center from 1979 to 1984. In the experiment, a round tensile specimen of sensitized Type 304 SS was being strained to failure under constant extension rate conditions in a dilute, oxygenated NaCl solution at 250 °C in an autoclave contained within a once-through flow loop. A crack had nucleated on the gauge section of the specimen, and the load had begun to fall, even though the strain continued to increase at a constant rate. While the load was falling, the pressure release valve failed and fluid was lost from the loop (overnight). The crack was found to have stopped growing, as witnessed by the increase in load. Qualitative estimates indicated that the cessation of crack growth occurred when the liquid level dropped below the level of the crack. Note that the steam pressure was such that capillary condensation should have maintained an environment within the crack. Since this early observation, the importance of the coupling between the internal and external environments has been confirmed by other experiments, particularly those of Manahan et al. [52], and Zhou and coworkers [53].

The role of the external environment is demonstrated by the calculated polarization data plotted in Fig. 11 [45]. Thus, it is seen that at low temperatures (e.g., 50 °C) significant polarization is predicted to occur in the external environment (170 mV). As the temperature increases, the polarization in the external environment is predicted to decrease, but, nevertheless remains significant at temperatures up to 250 °C. Even at higher temperatures, where the external polarization has been decreased to small values, owing to increased conductivity of the external environment and increased rate of oxygen reduction on the external surfaces, the external environment must still be considered, because of the need for charge conservation. The temptation has been to assume that, in this case, the potential in the solution at the crack mouth can be equated to $-E_{corr}$, and hence to eliminate the need to consider the external environment all together. However, this would be to assume that no potential gradient exists in the external environment, in which case no coupling current should be detected, again at odds with experiment [52].

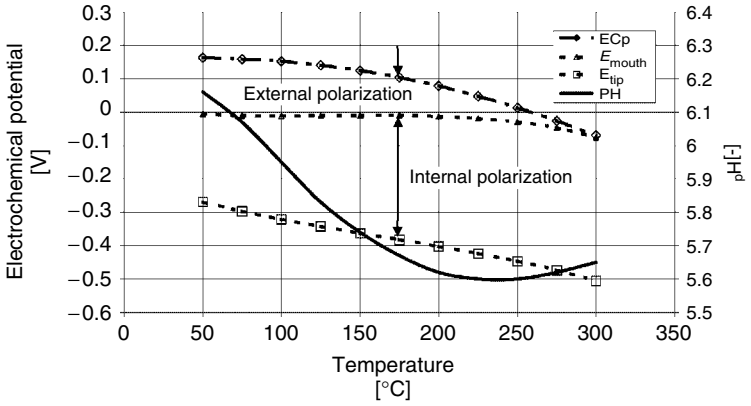


Fig. 11 The effect of temperature on the pH of the external environment, the electrochemical potential at the crack tip, E_{tip}, the potential at the crack mouth, E_{mouth}, and the potential in the external environment (ECP), during crack growth in Type 304 stainless steel in dilute sulfuric acid solution having an ambient temperature (25 °C) conductivity of 0.27 μS cm^{-1} and a dissolved oxygen concentration of 200 ppb. The data were calculated using the CEFM after calibration at 288 °C and assuming a crack-tip strain rate thermal activation energy of 100 kJ mol^{-1} (Congleton crack-tip strain rate model).

Finally, it is interesting to note that the decrease in the polarization in the external environment is predicted to be at the expense of an increase in the potential drop down the crack.

As noted elsewhere [43], the CEFM predicts that the CGR for given values of stress intensity, ECP, conductivity, and so on, depends upon the crack length. This prediction is consistent with available experimental data. Actually, in discussing this topic, it is necessary to differentiate between the mechanical crack length (MCL), which is traditionally referred to as the *crack length* in fractures studies, and the ''electrochemical crack length'' (ECL), as noted previously [54]. The ECL is defined as the path of least resistance through the solution from the crack front, where the coupling current is generated, to the external surface, where the current is consumed. For a CT specimen, the MCL corresponds physically to the distance between the load line (for example) and the crack tip, and this distance increases as the crack grows through the specimen. On the other hand, because of the through thickness nature of a crack in a CT specimen, the ECL is constant and is essentially independent of the MCL. Furthermore, a CT geometry yields a distribution in ECL depending upon where the current originates from at the crack front. This dependence is such that the CGR is highest at the crack edges (smallest ECL), but is lowest for the coupling current, which originates from the crack front in the center of the specimen (largest ECL) by virtue of the dependence of CGR on the ECL [54]. This phenomenon is responsible for the generation of convex crack fronts, when the crack grows by SCC, in contrast to the concave crack front obtained for creep crack growth [55]. An example of an extreme case of this phenomenon is shown in Fig. 12 [55], where it is evident that the crack propagated from the original fatigue

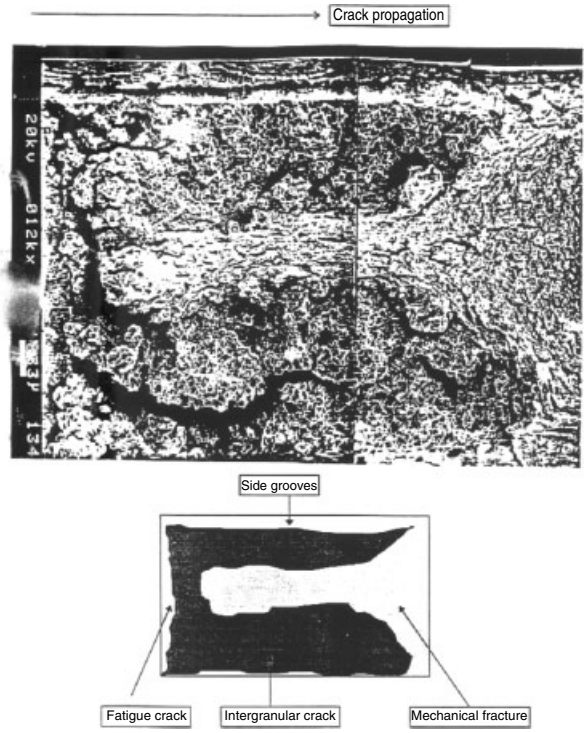


Fig. 12 Scanning electron micrograph of the fracture surface of sensitized Type 304 C(T) specimen after crack propagation in 15 ppm NaCl + 150 ppb O_2 at 250 °C under continuous stirring conditions [55].

crack along the side grooves, primarily because of the shorter ECL. Thus, while CT specimens are suitable for mechanical testing, their use for electrochemical SCC studies is questionable, a point that has been made previously [54, 55].

Calculated CGR versus ECP correlations for ECLs ranging from 0.1 to 50 cm and for a stress intensity factor of 27 MPa.$m^{1/2}$ and for the other, constant conditions, as stated in the figure, are shown in Fig. 13. In the low potential limit, no ECP dependence of the CGR is noted, because crack extension occurs only because of creep. At higher potentials, environmental effects become dominant such that, at sufficiently high potentials, the CGR increases in a roughly exponential manner with ECP. Note that the ECL is predicted to have a marked impact on the CGR: roughly 2 orders in magnitude for the ECL increasing from 0.1 to 50 cm. The practical importance of the dependence of CGR on the ECL has been previously stressed [54], in that it has a marked impact on the accumulated damage due to SCC.

Determination of the exact critical coupling current is best performed by plotting the coupling current versus CGR, as shown in Fig. 14. At the higher CGRs, the reader will see that the coupling current is of the order of 1–100 μA, which is consistent with experiment [52], noting that the value of the coupling current is sensitive

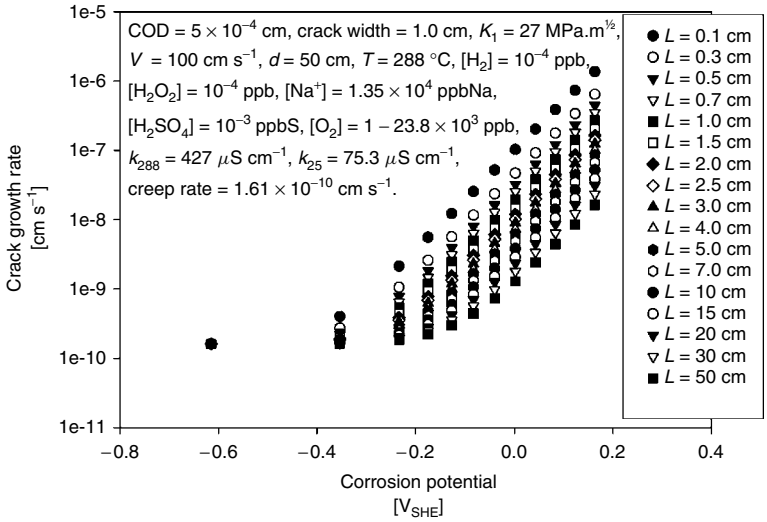


Fig. 13 Dependence of calculated crack growth rate on the electrochemical crack length for IGSCC in Type 304 SS in NaCl solution.

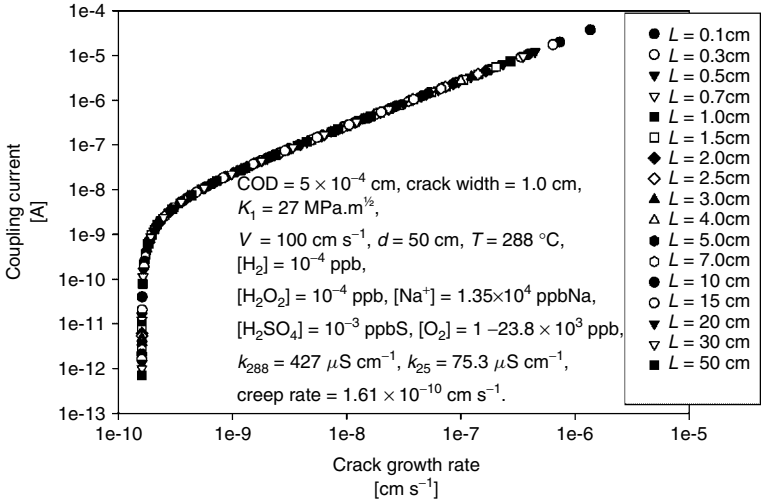


Fig. 14 Plots of coupling current versus crack growth rate for ECLs ranging from 0.1 to 50 cm for constant stress intensity and for the other, constant conditions stated in the figure [54].

to the electrocatalytic properties of the external surfaces (i.e., rate of reduction of oxygen). Interestingly, the form of the coupling current versus CGR relationship is independent of ECL: increasing ECL simply shifts the point for any given $[O_2]$ or ECP down the curve. This indicates that the critical coupling current is uniquely related to the CGR and, indeed, has been previously interpreted as being that

coupling current at which crack extension due to slip/dissolution/repassivation "outruns" that due to creep. In any event, Fig. 14 indicates that, for the conditions under consideration, the critical coupling current, as determined from the point of intersection of the extrapolated creep rate (lower left portion of the curve) and the extrapolated SCC rate (upper right portion of the curve), is about 3 nA for the particular specimen geometry considered.

The critical potential may now be determined by plotting the coupling current versus the corrosion potential, as shown in Fig. 15. The horizontal line, which represents the critical coupling current (3 nA), intersects the coupling current versus ECP correlations at potentials ranging from $-0.40\ V_{SHE}$ for an ECL of 0.1 to $-0.15\ V_{SHE}$ for an ECL of 50 cm. The origin of the dependence of CGR on ECL is that, as the ECL increases, the potential drop down the crack increases correspondingly, and hence less potential drop is available in the external environment to drive the oxygen reduction reaction. Because the crack cannot grow any faster than the oxygen reduction reaction can consume the coupling current, the coupling current must decrease (see Fig. 15) and, hence, so must the CGR. If the oxygen reduction reaction on the external surface is catalyzed (e.g., by deposition an electrocatalyst, such as Pt), the coupling current is increased, as observed experimentally [52], and so must be the CGR, except if the ECP is reduced, for any reason.

An important finding of this previous work [54] is that the critical potential for IGSCC is predicted to depend upon the ECL. Thus, with reference to Fig. 15, it is evident that E_{IGSCC} increases (becomes more positive) as the ECL increases, changing from $-0.40\ V_{SHE}$ to $-0.15\ V_{SHE}$ as the ECL increases from

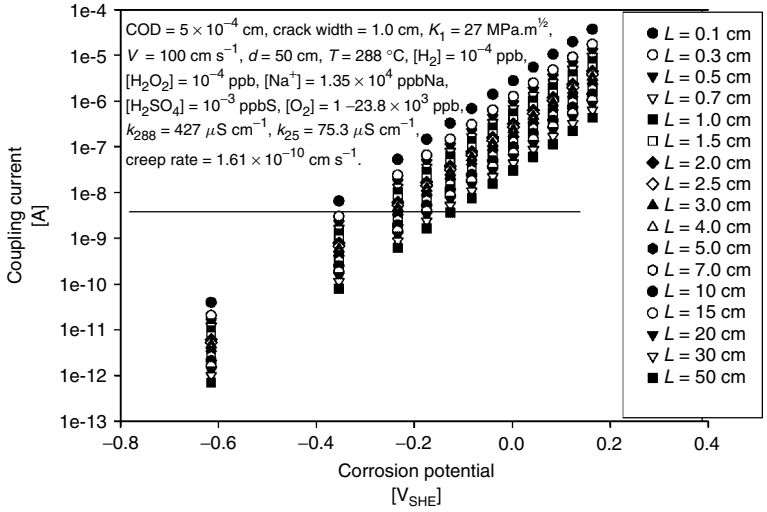


Fig. 15 Plots of coupling current versus ECP (corrosion potential) for a range of ECLs extending from 0.1 to 50 cm for the conditions shown in the figure. The horizontal line corresponds to the calculated critical coupling current, as determined from Fig. 14, which defines the critical potential, E_{IGSCC} for stress corrosion cracking [54].

0.1 to 50 cm, respectively, as noted above. The importance of this prediction is that it suggests a mechanism by which cracks will cease to propagate after a certain ECL is attained. Again with reference to Fig. 15, we note that if the corrosion potential of a component was $-0.20\ V_{SHE}$, SCC (EAC) should cease after an ECL of 3.0 cm is attained, because at that ECL the critical potential becomes more positive than the corrosion potential [54]. The reader will note that the predicted critical potential for IGSCC for the shortest ECL considered (0.1 cm), $-0.40\ V_{SHE}$, is significantly more negative than the often-quoted value of $-0.23\ V_{SHE}$ for BWR coolant conditions [3]. Notwithstanding that the ECL corresponding to the BWR

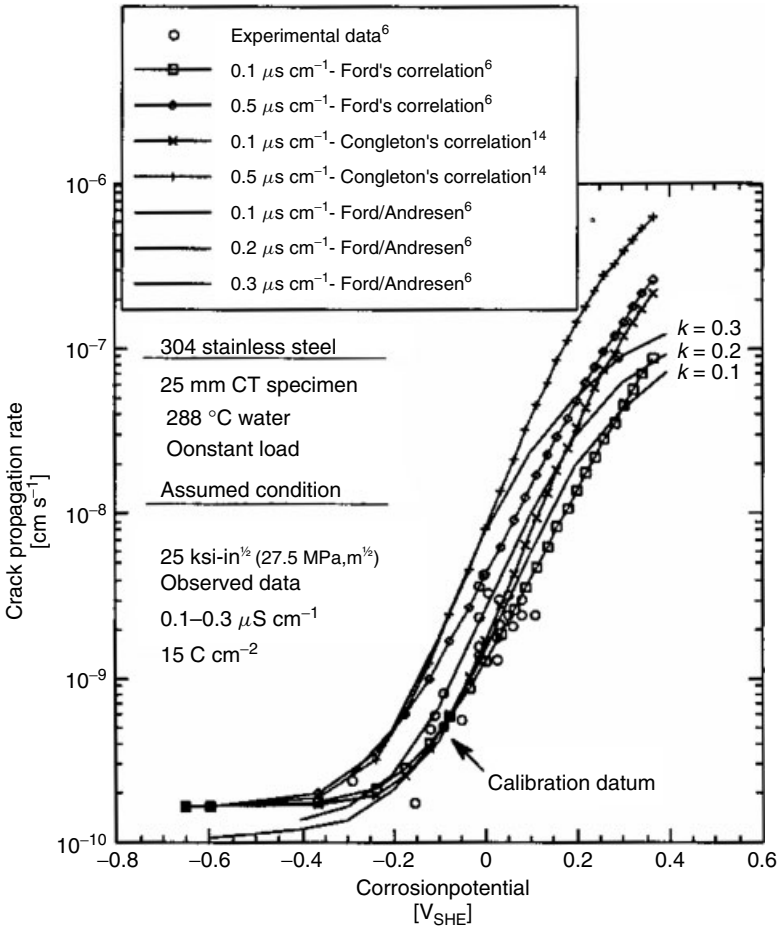


Fig. 16 IGSCC crack growth rate versus corrosion potential (ECP) for sensitized Type 304 SS in high-temperature water as a function of solution conductivity (κ_{25}) for given values for the degree of sensitization (15 C cm^{-2}) and stress intensity (27.5 MPa.$m^{1/2}$). The experimental data were measured for ambient temperature conductivities, κ_{25}, ranging from 0.1 to 0.3 μS cm^{-1}. The citations given in the figure are those presented in Ref. 43.

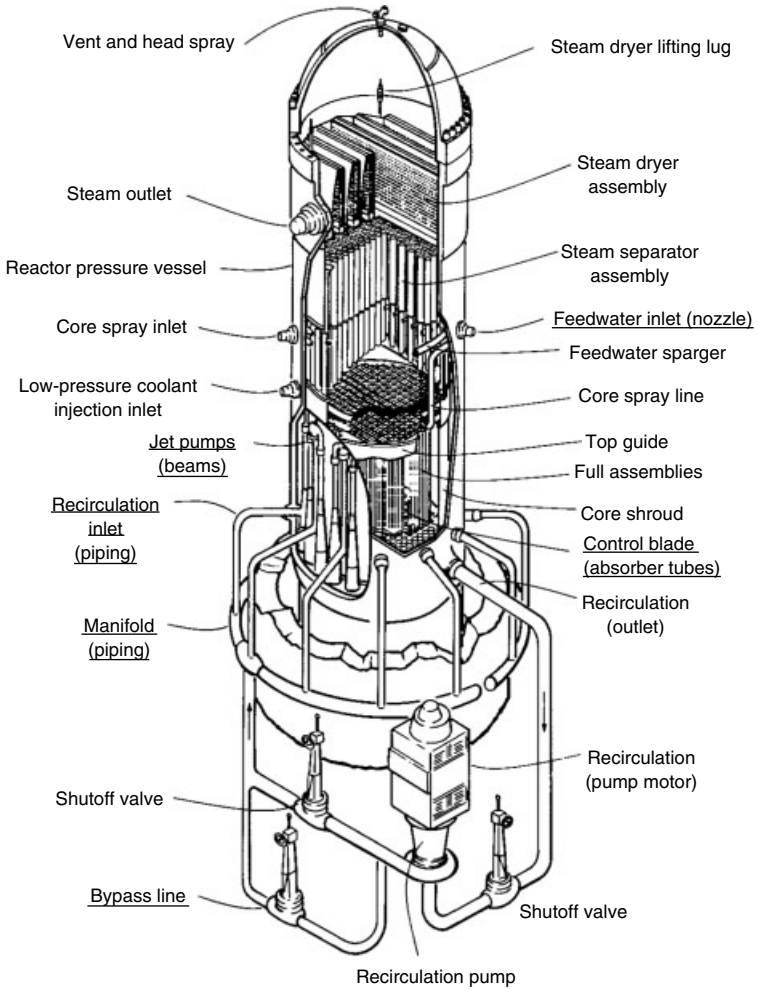
case is unknown, the critical potential was determined from constant extension rate tests (CERTs) at a finite extension rate, the conductivity for the case calculated is significantly higher than is a BWR coolant, and the critical potential is known to shift in the negative direction with increasing ionic strength of the solution [56], the fact that the calculated critical potential is 170 mV more negative than for the BWR case for the same temperature is not surprising.

That a critical potential for IGSCC exists is clearly shown in Fig. 16, which compiles experimental CGR versus ECP data taken from a variety of sources together with values calculated using the original CEFM [43] using different crack-tip strain rate models due to Andresen and Ford [40] and Congelton [57]. The model due to Andresen and Ford is essentially empirical and yields a crack-tip strain rate that increases with the fourth power of the stress intensity factor. A satisfactory mechanistic explanation for this dependence has never been advanced, to the knowledge of the author. On the other hand, the crack-tip strain rate model of Congelton is based on linear elastic fracture mechanics (LEFM) and hence is well founded in physical theory. In any event, the calculated correlations shown in Fig. 16 were generated by the CEFM using both crack-tip strain rate models after calibration of the model on a single CGR versus ECP, conductivity, stress intensity, degree of sensitization (DOS), and temperature datum, as indicated in the figure. The creep CGR at very negative potentials corresponds to that calculated from the Wilkinson and Vitek model [51], as described above. At high potentials, the CGR is found and predicted to increase almost exponentially with potential; it is in this region that electrochemical effects dominate. The critical potential corresponds to the condition where the creep CGR equals the electrochemical CGR, with this condition occurring over the potential range of -0.2 to $-0.4\ V_{SHE}$, which is in excellent agreement with the theoretically estimated values shown in Fig. 16.

9.5 Electrochemistry of Boiling Water Reactors

As noted above, the increasing incidence of IGSCC in sensitized stainless steel components in BWR HTCs has led to the development of radiolysis models for calculating the concentrations of electroactive species, such as H_2, O_2, and H_2O_2, as a function of the reactor operating parameters and the concentration of hydrogen added to the feedwater [6–30]. Because these species are electroactive, they are instrumental in establishing the ECPs of components within the HTC [6, 9, 13, 14, 20–30]. Extensive work in many laboratories worldwide has established that sensitized Type 304 SS becomes increasingly susceptible to IGSCC in high-temperature aqueous solutions as the ECP is increased above a critical value [10].

A cut-away schematic of a BWR equipped with external coolant pumps is shown in Fig. 17. The recirculation system comprises the external piping, pumps, and valves located at the lower region of the vessel. The penetrations through the bottom of the vessel that contain the control rod drive mechanisms that are vital for controlling the reactivity of the core are not shown in Fig. 17. Of particular importance, as far as the integrity of the reactor is concerned, is the control rod drive tubes and related mechanisms, because they are in contact with the coolant and



Key components of the BWR. the parts names in red indicates areas of IGSCC.

Fig. 17 Cut-away schematic of a BWR vessel and associated piping system for a BWR equipped with external coolant pumps. The underlined labels are indicating regions of IGSCC. The control rod drive mechanisms and penetrations that are located at the bottom of the vessel are not shown.

hence are susceptible to SCC. However, as indicated in Fig. 17, cracking has occurred in a variety of vessel internals, including the control rod drive tubes, so there is nothing peculiar about that particular location. For example, almost complete, circumferential cracking of the core shroud adjacent to a weld in the upper regions was detected in one BWR in the United States. In this particular case, the question begged is, How (or) did the crack grow so rapidly to almost completely sever the shroud? Given that the shroud is about 1800 cm in circumference, and noting that the ECP

at the inner surface of the shroud is of the order of 0.3 V_{SHE}, the CGR should be about 10^{-7} cm s^{-1} (Fig. 10). Thus, the time over which the crack must have grown is 1.8×10^{10} seconds or 571 years! Clearly, the damage was not caused by a single crack. The most likely scenario is that multiple cracks nucleated on the inside of the shroud and coalesced to form a single crack. If this is so, then the cracks must have nucleated in similar times, in order not to have left multiple ligaments, and hence the nucleation process must have been deterministic, rather than stochastic, in nature.

The primary coolant circuit of a BWR equipped with external recirculation pumps is shown schematically in Fig. 18. This design, which is typical of late model BWRs supplied by the General Electric Company, has a history of IGSCC in sensitized, HAZs adjacent to welds, as discussed above, and that, as time passed from when cracking was first detected in the early 1970s, cracking has progressively spread from the small-diameter pipes in the recirculation system to larger diameter pipes in the same system and eventually to in-vessel components, where safety issues become of paramount importance. In a BWR, the primary coolant circuit contains liquid water in all sections, except in the fuel channels, where boiling is initiated, and in the upper plenum, where the steam is separated from the water, with the former being dried and transferred to the turbines and the latter being mixed with feedwater in the mixing plenum and returned to the primary circuit. Without any doubt, a primary cause was the use of Type 304 SS, which is easily sensitized upon welding and hence is susceptible to IGSCC under BWR coolant conditions. Tacit recognition that IGSCC might be a problem is evidenced by the fact that the stabilized grade, Type 347 SS, was specified for use in BWRs in Germany as early as the 1960s. Type 347 SS contains a small amount of niobium to preferentially react with carbon and hence to reduce the

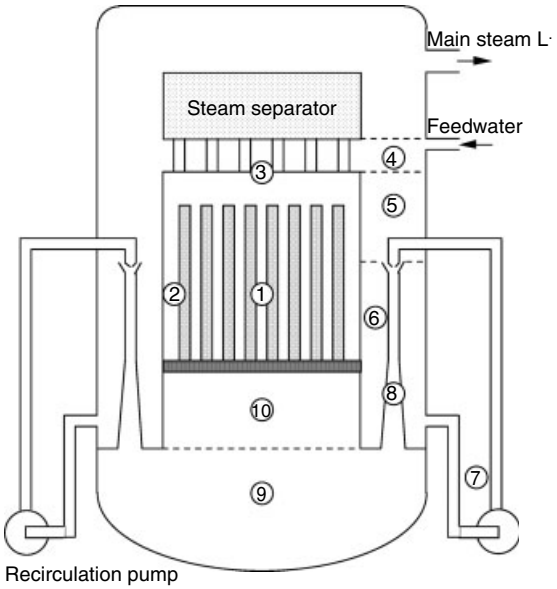
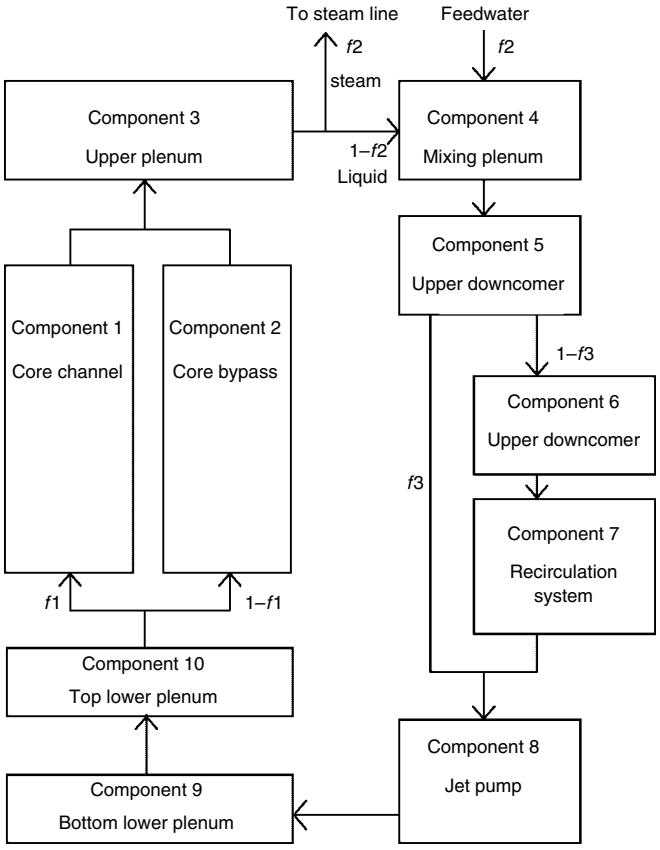


Fig. 18 Schematic of the primary coolant circuit of a BWR having external pumps. The regions are identified as follows: (1) core channels, (2) core bypass, (3) upper plenum, (4) mixing plenum, (5) upper downcomer, (6) lower downcomer, (7) recirculation system, (8) jet pump, (9) bottom of lower plenum, (10) top of lower plenum.

activity of free carbon in the matrix, to the point that the formation of chromium carbides (e.g., $Cr_{23}C_7$) is greatly inhibited or does not occur at all. It is the formation of chromium carbides on the grain boundaries within the HAZs that result in chromium depletion adjacent to the boundaries and hence renders them susceptible to intergranular attack (including IGSCC). However, both laboratory work and field observation have demonstrated that, given the presence of susceptible microstructure and tensile loads, the value of the ECP, and hence the electrochemistry of the coolant, is paramount in determining the occurrence of IGSCC.

In modeling the thermal hydraulics of a BWR, the flow scheme shown schematically in Fig. 19 is commonly invoked. This diagram indicates three bifurcations



*f*1 : Fraction of mass flow into core channel
*f*2 : Fraction of mass flow out of the core circuit or
Ration of feedwater mass flow to total mass flow
*f*3 : Fraction of mass flow from upper downcomer into jet pump

Fig. 19 Flow path of coolant in the primary coolant circuit of a BWR having external coolant pumps. The flow fractions for each region are labeled f1, f2, and f3.

in the flow: at the bottom of the core into the fuel channels (boiling, 90%) and into the core bypass (nonboiling, 10%); in the upper plenum with the separation of steam (15%) from the water that is transferred to the mixing plenum (85%); and in the upper downcomer, where part of the flow is directed through the recirculation system (30%) with the balance being sucked into the jet pumps and injected into the bottom of the lower plenum (70%), with the percentages in parentheses being approximate values. The steam, after passing through the turbines, is condensed and returned to the core as feedwater. Radiolysis occurs only within the core, core bypass, and in the downcomer with the most intense radiolysis taking place in the fuel channels and much less occurring in the bypass and downcomer, because of the lower dose rates of γ photons and neutrons. Typical dose rates in the fuel channels are 8×10^4 rad s^{-1} (5×10^{18} eV cm^{-3} s^{-1}) and 4×10^5 rad s^{-1} (2.5×10^{19} eV cm^{-3} s^{-1}), respectively. These dose rates are considerably smaller than those present in a PWR core (see below), corresponding to a higher energy-density core in the case of the latter. Prior to appreciation of the role of electrochemistry in the cracking of primary coolant circuit components, the philosophy with regard to BWR coolant chemistry had been to maintain it as pure as possible; thus the Electric Power Research Institute instituted the ''EPRI Guidelines'' on water chemistry, in which progressively tighter limits were placed on the concentration of chloride ion, for example. While it was recognized that the concentration of oxygen in the coolant might, somehow, be connected to the cracking problem, the production of oxidizing conditions via the radiolysis of water was not generally appreciated. Indeed, some utilities took measures to ''deaerate'' the coolant prior to start-up on the belief that reducing the concentration of oxygen established when the system was open to the atmosphere during refueling would reduce the susceptibility to IGSCC. The strategy failed. At the same time (late 1970s and early 1980s), work in a number of laboratories (e.g., Ohio State University [56]) demonstrated that the electrode potential was an important parameter in determining the susceptibility of sensitized, austenitic stainless steels to IGSCC in high-temperature aqueous systems. Attention was then focused on the need to reduce the high ECP that is established by the buildup of radiolytically produced, oxidizing species, such as O_2 and H_2O_2, recognizing that hydrogen was preferentially stripped from the coolant by boiling in the fuel channels. This, in turn, led to the introduction, by the General Electric Company and EPRI, of HWC, which had been pioneered in Sweden, in which small amounts of hydrogen are added to the feedwater to suppress the ECP.

Since its first introduction in the early 1990s, HWC [58, 59] has undergone steady development, particularly with regard to enhancing the effectiveness of the technique by noble metal chemical additions (NMCA), after the limitations of HWC had been made clear by electrochemical analysis. By the mid-1990s, it was evident that many of the answers that were sought with regard to the application of HWC would be best provided by sophisticated BWR coolant chemistry models [6, 9, 13, 14, 19, 26, 27]. Importantly, these models calculate the parameters of prime importance; the ECP and the resulting CGR. Accordingly, because of the limitations of space, the remainder of this discussion will be devoted to the models developed by the author.

These codes were based on the radiolysis model originally proposed by Burns and Moore [6], comprising the first 34 reactions, together with a reaction describing the thermal decomposition of hydrogen peroxide ($H_2O_2 = H_2O + 1/2O_2$). Indeed, the model was calibrated to reactor recirculation oxygen levels by adjusting the rate constant for this single reaction; the rate constant has never needed to be adjusted after the first calibration on data from the Duane-Arnold plant in the United States and the value of the adjusted rate constant was within the range reported from laboratory studies [6]. With this single calibration, the original code developed by the author and his colleagues, DAMAGE-PREDICTOR, and its later derivatives have been found to accurately describe the water chemistry of BWRs across the entire spectrum with regard to plant response to hydrogen added to the reactor feedwater.

A unique feature of DAMAGE-PREDICTOR and its derivatives that is not present in other codes, except for that of Yeh [20–29, 1] (Dr Tsuang Kuang Yeh developed DAMAGE-PREDICTOR in the author's group in partial fulfillment of the requirements for a Ph.D. in Nuclear Engineering at the Pennsylvania State University from 1992 to 1996 and has continued to develop radiolysis codes for BWRs since returning to Taiwan.), is the ability to calculate the ECP and the CGR using the MPM and the CGR using the CEFM described previously. Derivatives of DAMAGE-PREDICTOR developed in the author's group, including ALERT, REMAIN, and FOCUS, also calculate the IGSCC damage integrated over the CEP (operational history of the reactor). These four codes incorporate deterministic modules for estimating the specie concentrations, the ECP, and CGR of stainless steel components at closely spaced points around the coolant circuit, as a function of coolant pathway geometry, reactor operating parameters (power level, flow velocity, dose rates, etc.), coolant conductivity, and the concentration of hydrogen added to the feedwater. DAMAGE-PREDICTOR, which has been used to model 14 BWRs worldwide, has been validated by direct comparison with plant data (e.g., at the Leibstadt BWR in Switzerland), and is found to accurately simulate HWC (Figs 5 and 20). The code has also been used to explore various enhanced versions of HWC and completely new strategies, such as those that employ noble metal coatings and dielectric coatings, respectively. Two of the component models of DAMAGE-PREDICTOR, in fact, *predicted* quantitatively the effectiveness of dielectric coatings for inhibiting crack growth in stainless steels in high-temperature water, and these predictions have been validated by direct experiment [53]. The application of these codes in modeling the accumulation of damage in BWRs has been described at length in the literature [6, 9, 13, 14, 19, 20], so only a summary of the more important results will be given here.

The process of predicting the accumulation of IGSCC damage in sensitized Type 304 SS in a BWR coolant circuit begins with specifying the CEP, which is an $N \times M$ matrix containing $M \times$ (time (t) and $N - 1$ property entries) with each set defining a single state point of reactor operation. The $N - 1$ properties include the following: (1) distance from bottom of core (cm), (2) temperature (°C), (3) flow velocity (cm s^{-1}), (4) steam velocity (cm s^{-1}), (5) void fraction, (6) hydrodynamic diameter of the channel (cm), (7) neutron dose rate (R s^{-1}), (8) γ photon dose rate R s^{-1}),

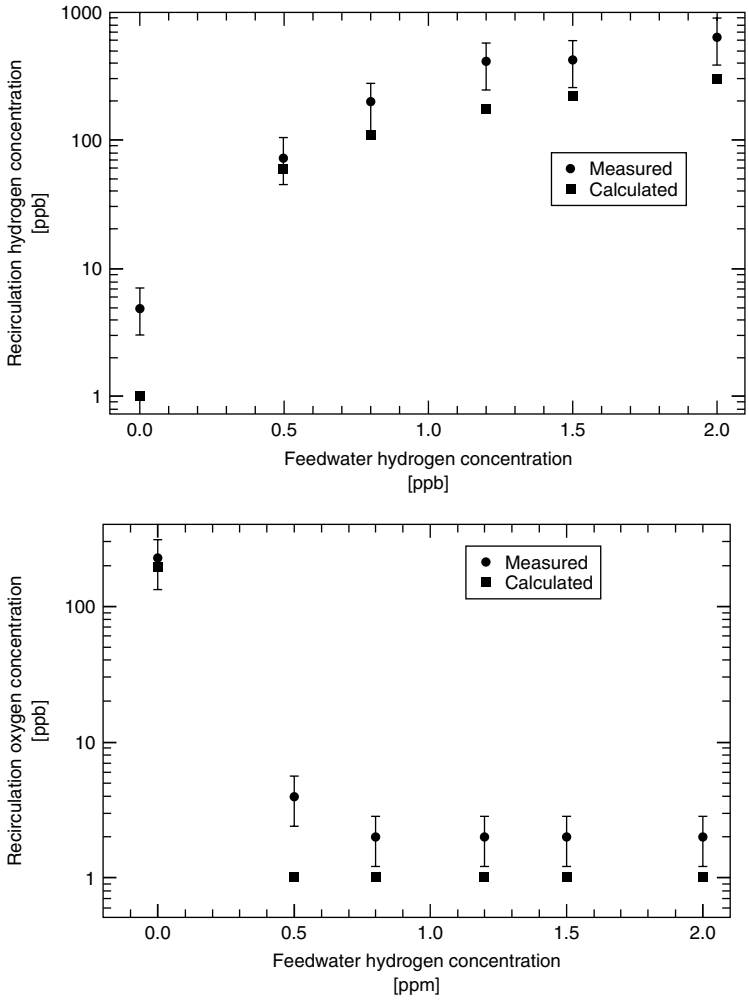


Fig. 20 Comparison of measured and calculated $[H_2]$ and $[O_2]$ in the recirculation piping system from a "double-blind" HWC "mini-test" at the Leibstadt BWR [19].

(9) catalysis factor for the decomposition of H_2O_2, (10) exchange current density multiplier for the HER, (11) exchange current density multiplier for the OER, (12) exchange current density multiplier for the HPER, (13) initial stress intensity for a 0.5-cm crack ($MPa.m^{1/2}$), and (14) conductivity at ambient temperature ($S\ cm^{-1}$). The exchange current density multipliers are included, so that catalytic protocols, such as NMCA/HWC, and inhibitive strategies, such as the use of dielectric coatings (see below), can be modeled. The mixing table that defines how the flow is distributed among the components around the circuit is also included (Fig. 19). A table of this type is prepared for each region of the reactor, as defined in

Fig. 20, and for each state point (position on the CEP). Given that the CEP might be defined by more than 100 state points, the input is voluminous indeed. Other reactor details are included in ancillary tables.

The output contains position and temperature data, the concentrations of various species, the ECP and redox potential (potential measured on platinum), the CGR, the conductivity at ambient temperature, and the conductivity at the operating temperature. As noted previously, the ECP is calculated from the concentrations of O_2, H_2O_2, and H_2, with these three species dominating the mixed potential. In the case of the redox

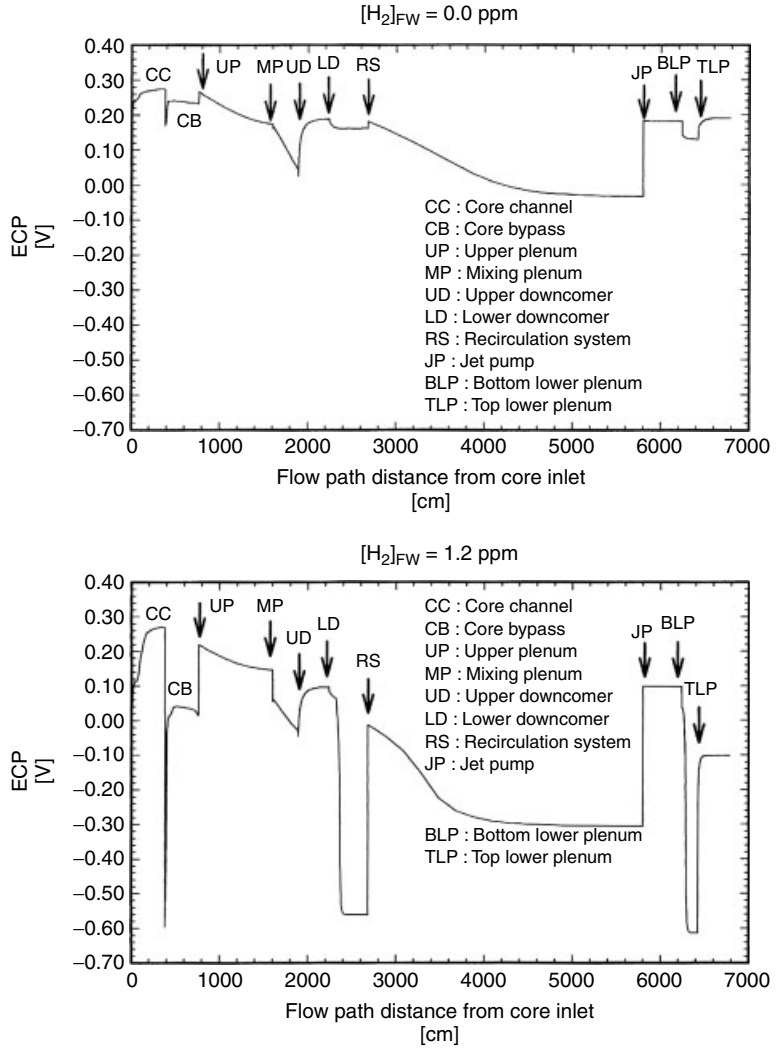


Fig. 21 Predicted corrosion potential (ECP) versus flow path distance from the bottom of the core for 0 and 1.2 ppm of hydrogen added to the feedwater of the Leibstadt BWR [19]. The locations correspond to those given in Fig. 19.

potential, the anodic partial current is set equal to zero and the exchange current density multipliers are set equal to values that are appropriate for platinum.

Typical predictions of ECP and CGR around the primary coolant circuit of a GE Model 6 BWR (Leibstadt) for a particular state point, as a function of hydrogen added to the feedwater, are shown in Figs 21 and 22. The two feedwater hydrogen concentrations chosen (0 and 1.2 ppm) correspond to NWC and ''super'' HWC, respectively. The CGR plots are for a ''standard'', 0.5-cm crack loaded to a stress

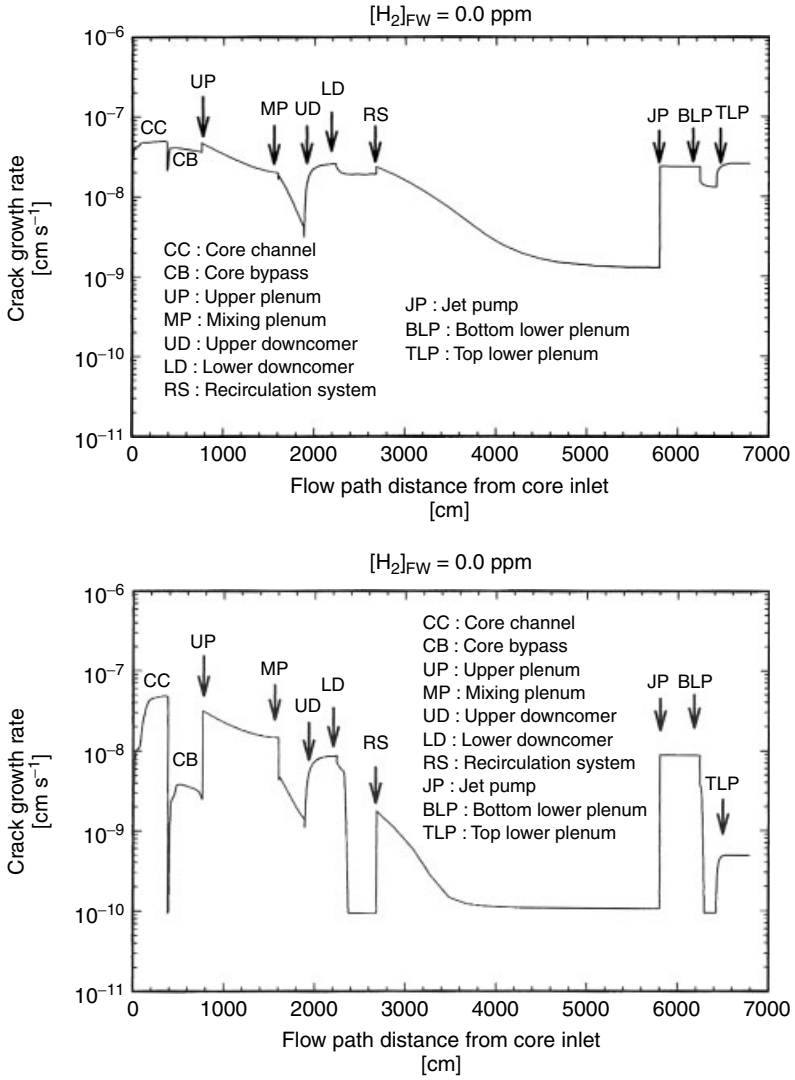


Fig. 22 Predicted crack growth rate versus flow path distance from the bottom of the core for 0 and 1.2 ppm of hydrogen added to the feedwater of the Leibstadt BWR [19].

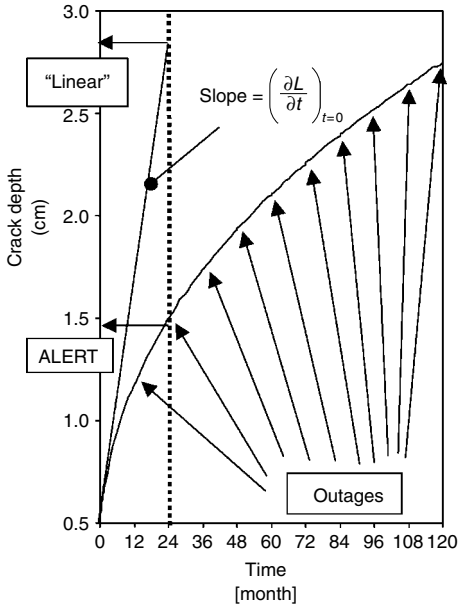


Fig. 23 Illustration of the nonlinearity of the accumulated damage and inadequacy of the linear approach for extrapolating crack propagation for future operating periods.

intensity factor of 27.5 MPa.m$^{1/2}$ and for an ambient temperature conductivity of 0.1 μS cm^{-1}, as a function of feedwater hydrogen concentration. The ECP and CGR are related to the concentrations of H_2, O_2, and H_2O_2 in a rather complex manner, in addition to depending on flow rate and temperature [43, 58]. These complex relationships cannot be captured by empirical methods, simply because the responses of the ECP and CGR to each of the independent variables, and each combination of variables, are highly nonlinear. It is for this reason that the MPM and the CEFM were originally developed. Note that the critical potential for IGSCC, as recommended by the NRC, is $-0.23\ V_{SHE}$. One sees from this analysis that only a few regions in the primary coolant circuit of this reactor are predicted to be protected by HWC. Importantly, however, much of the downcomer and the lower plenum are predicted to be protected by the addition of modest amounts of hydrogen to the feedwater.

A somewhat less pessimistic picture is obtained by recognizing that a CGR of $<5 \times 10^{-9}$ cm s^{-1} for a 0.5-cm crack loaded to $K_I = 27.5$ Mpa.m$^{1/2}$ is probably tolerable (≈1.6 mm year^{-1}), given that the nucleation time for cracks is reckoned to be quite long (years) and that the CGR is predicted to decrease with increasing crack length. By this criterion, HWC appears to be considerably more effective at protecting the primary coolant circuit of a BWR than when judged upon the basis of the ECP.

As noted above, DAMAGE-PREDICTOR, ALERT, and REMAIN contain the CEFM for predicting CGR. The CEFM has been extensively tested against experiment, as discussed elsewhere [43, 45, 46] and earlier in this chapter. The latest code, FOCUS, contains even more advanced versions of the MPM and CEFM, where the latter allows the CGR to be calculated over a much wider range of conductivity and incorporates the thermally activated crack-tip strain rate, as discussed earlier in this chapter. The deterministic nature of the CEFM implies that minimal calibration is required. Accordingly, because it captures vital relationships between the CGR and various independent variables, it can be used to model regions in a reactor for which insufficient data exist for the reliable calibration of empirical models. For example, the CEFM yields the CGR as a function of crack length. This relationship, which is not captured by any empirical model, to the author's knowledge, is essential for the prediction of integrated damage (i.e., crack length as a function of time for a proposed

operating history), because the CGR decreases as the crack length increases [54], as shown in Fig. 23. This is due to an increase in the potential drop down the crack, even though the mechanical driving force (the stress intensity) is maintained constant. The calculations shown in this figure were carried out on a crack on the inside of the core shroud adjacent to the H-3 weld near the top of the shroud over a 120-month operating history of a reactor in service. If the dependence of CGR on crack length is not recognized, the integrated damage function (IDF) is overpredicted by several hundred percent, thereby leading to a much more pessimistic evolution of damage (Fig. 23).

Application of the models at each state point yields a CGR, but the CGR depends upon the value of parameters along the CEP and upon the stress intensity (although only weakly) and the crack length. Thus, beginning at the initial point on the CEP, the crack length is updated successively for the next state point. Thus, the crack length, x_N, over the anticipated service time of a component, T, is obtained by an accumulation of the crack advances over N periods of time $\Delta t_1, \ldots, \Delta t_i, \ldots \Delta t_N$ [41].

$$L_i = L_{i-1} + CGR_i \cdot \Delta t_i, i = 1, \ldots, N \tag{27}$$

with

$$T = \sum_{i=1}^{N} \Delta t_i \tag{28}$$

and

$$K_{I,i} = \sigma(t) Y \sqrt{L_i} \tag{29}$$

where σ (t) is the (time dependent stress (applied and residual), Y is a factor that depends on crack geometry and can be generally obtained from tables of stress

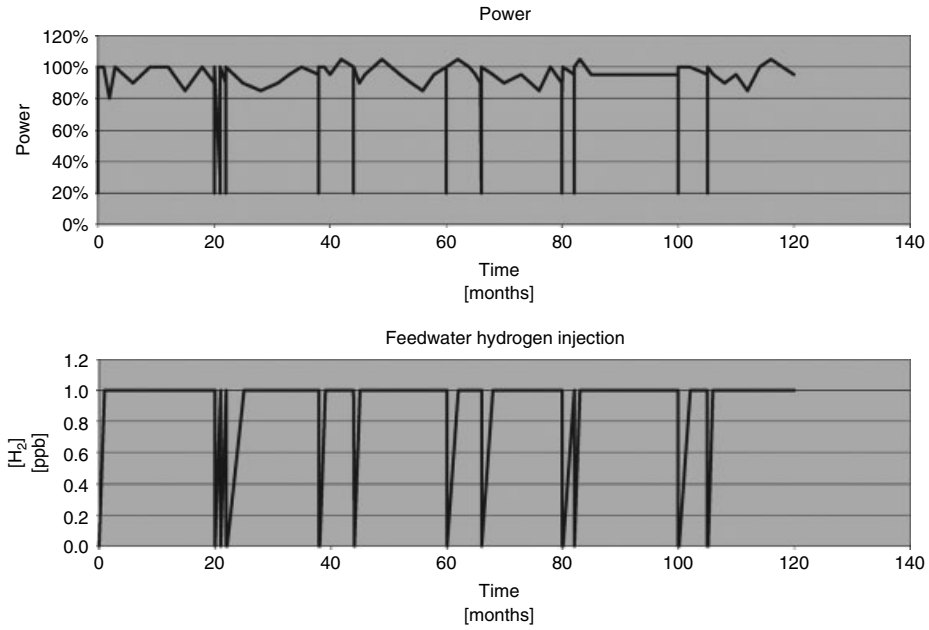


Fig. 24 Corrosion evolutionary path (CEP) defined in terms of the prime variables of reactor power and hydrogen added to the reactor feedwater.

intensity factors, and the CGR, CGR_i, is presumed to be time-independent for each interval, Δt_i, in that it depends on the crack length (through the stress intensity, K_I, and because of changes in the current and potential distributions in the crack-internal and crack-external environments). The initial crack length, L_0, corresponds to the depth of a preexisting crack (as may have been detected during an inspection or assumed for a safety analysis scenario).

A series of ''strip charts'' is presented illustrating the modeling of the extension of a crack adjacent to a weld on the inner surface of the shroud of a GE Model 6 BWR as the reactor traverses the CEP, in this case defined in terms of reactor power as the primary variable (Fig. 24). In the second chart (Fig. 25), the feedwater hydrogen concentration is also plotted, with the decreases corresponding to reactor shutdown for either refueling or for unscheduled reasons. During these shutdowns, the conductivity of the coolant changes owing to changing temperature and the release of ions from crevices and deposits (''hideout return''), and other parameters change as well, with the changes either being introduced in ancillary tables or calculated internally in the code. The CEP depicted in Fig. 24 is arbitrarily chosen, but is considered to be quite realistic.

The stress intensity is calculated on the assumption that the stress on the crack does not change with time, that is, the increase in K_I is due entirely to the increase in crack length, as indicated by Eq. (29). The reader will also note that the CGR increases significantly during shutdown and start-up (Fig. 26); this is a consequence of the parabolic variation of CGR versus temperature, as shown in Fig. 10, the increase in conductivity of the coolant, and the reduction in hydrogen concentration in the coolant. Note that the ECP also undergoes positive excursions during shutdown/start-up.

The integral damage function, that is, the crack length, versus time of reactor operation is shown in Fig. 26. As expected, the crack length increases more or less monotonically with time, with perturbations that are due to chemical and electrochemical upsets during changes in power level (including shutdowns and start-ups), in a manner that reflects increasing IR potential loss down the crack, as previously noted.

The models (DAMAGE-PREDICTOR, ALERT, REMAIN, and FOCUS) permit modeling of the accumulated integral damage (crack length) for different operating scenarios. Three scenarios are shown in Fig. 27: (1) normal water chemistry, (2) hydrogen water chemistry (1 ppm H_2 in the reactor feedwater, (3) NWC for 5 years followed by HWC for the remaining 5 years. These calculations correspond to the CEP defined in Fig. 24. Under NWC, the crack is predicted to grow from an initial length of 0.5 cm to about 2.7 cm after 10 years of operation. If the reactor had operated continuously on HWC, however, the predicted crack length after 10 years of operation is about 1.1 cm, an extension of only 0.6 cm compared with 2.2 cm under NWC. If HWC were adopted after 5 years of operation, the predicted crack depth is 2.2 cm over the 10-year composite period, corresponding to a crack extension of 1.7 cm. In this case, the benefits of HWC are not so apparent, because most of the crack extension is predicted to have occurred over the first five years when the crack length was relatively small. Clearly, benefits accrue to the early implementation of HWC.

The codes DAMAGE-PREDICTOR, ALERT, REMAIN, and FOCUS contain

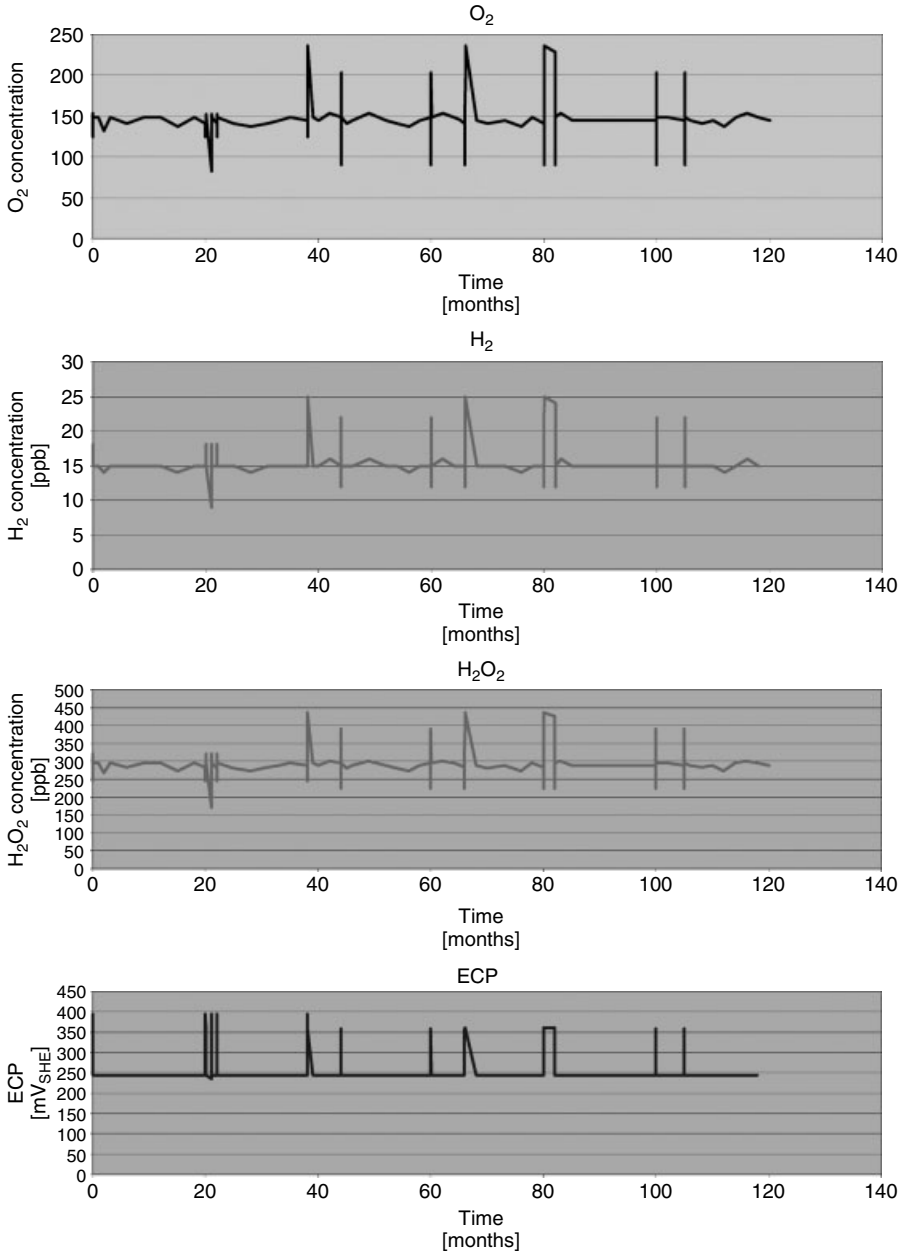


Fig. 25 Plots of the concentrations of oxygen (ppb), hydrogen (ppb), and hydrogen peroxide (ppb), and the calculated ECP (mV_{SHE}) at the top inner surface of core shroud as a function of time along the CEP. Note that the concentration of hydrogen is considerably lower than the feedwater hydrogen, because of the mixing of flows from the lower plenum and because hydrogen is stripped from the coolant by boiling in the core.

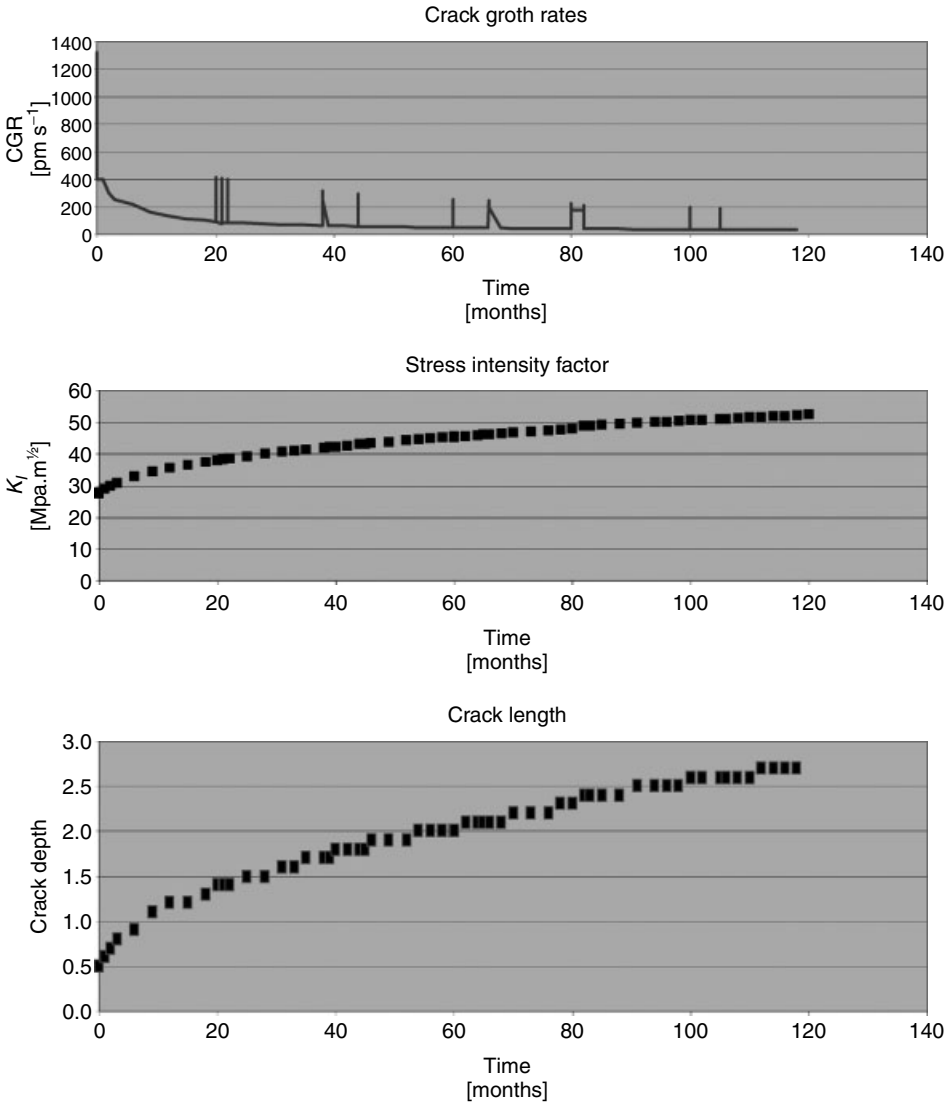


Fig. 26 Plots of the concentrations of crack growth rate (pm s^{-1}), stress intensity (MPa.$m^{1/2}$), and the crack length (mm) at the top inner surface of core shroud as a function of time along the CEP.

the necessary facilities for modeling enhanced hydrogen water chemistry (EHWC), as affected by the use of catalytic coatings (i.e., noble metal coatings, NMCA), and other advanced remedial measures, such as surface modification by dielectric coatings (SMDC) and ultra low conductivity operation (ULCO). A considerable achievement of the MPM, which is incorporated in the above codes) was the prediction that dielectric coatings represented a viable, and indeed an advantageous, alternative to noble metal coatings – a prediction that was made in

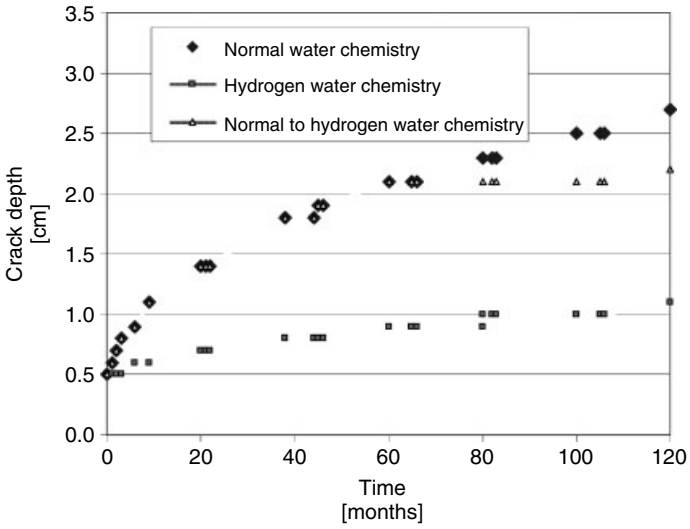


Fig. 27 Integrated damage functions (crack length) versus reactor operating time for three operating scenarios: (1) normal water chemistry, (2) hydrogen water chemistry (1 ppm H_2 in the reactor feedwater, (3) NWC for 5 years followed by HWC for the remaining 5 years [19].

the early 1990s and has been confirmed experimentally [53].

The effectiveness of both strategies arise from modification of the exchange current densities for the redox reactions (oxidation of hydrogen and the reduction of oxygen and hydrogen peroxide) that occur on the steel surface [9, 14, 26, 29, 43]. In the case of the noble metal coatings, the exchange current densities are increased, with the greatest impact being on the HER, but only if hydrogen is present in stoichiometric excess. This renders hydrogen to be a much more effective reducing agent than it is in the absence of the noble metal, thereby making it much more effective in displacing the ECP in the negative direction. In the case of dielectric coatings, the lower exchange current densities render the metal less susceptible to the ECP-raising oxidizing species, with the result that the ECP is displaced in the negative direction, even in the absence of hydrogen added to the feedwater. The multiplication factors that have been obtained experimentally are considerably smaller than the 10^4 assumed above [53], but they are, nevertheless, sufficiently large to indicate a significant impact of noble metal coatings on the effectiveness of HWC.

The calculated ECP and CGR values plotted in Figs 28 and 29, respectively, for general catalysis, in which the exchange current densities are increased arbitrarily by a factor of 10^4, show that this strategy is highly effective in displacing the ECP in the negative direction when a stoichiometric excess of hydrogen is present in the coolant, as noted above. If excess hydrogen is not present, as in the case of NWC, the ECP is predicted to be shifted in the positive direction, in agreement with the laboratory measurements of Yeh et al. [26, 29]. The CGR is predicted to respond accordingly by becoming much lower under

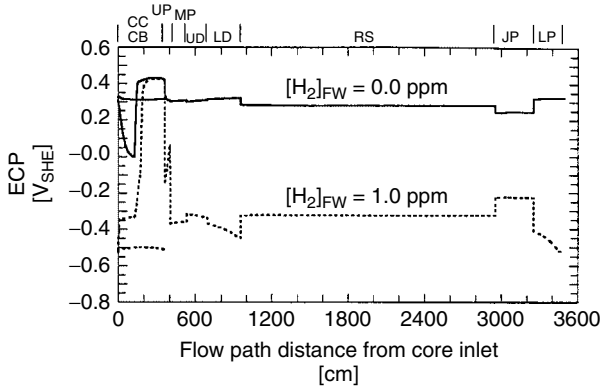


Fig. 28 ECP variation in the heat transport circuit of Dresden-2 employing general catalysis (exchange current density multiplier 10^4) under NWC ($[H_2]_{FW} = 0.0$ ppm) and HWC ($[H_2]_{FW} = 1.0$ ppm).

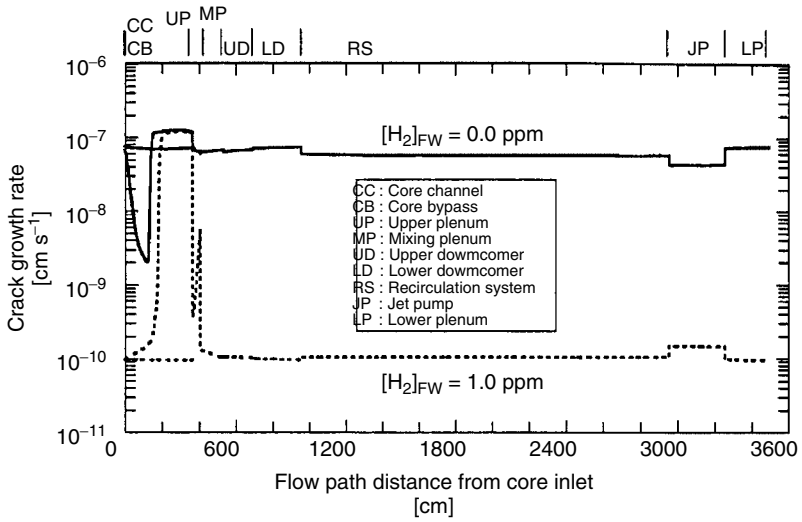


Fig. 29 Crack growth rate variation in the heat transport circuit of Dresden-2 employing general catalysis (exchange current density multiplier 10^4) under NWC ($[H_2]_{FW} = 0.0$ ppm) and HWC ($[H_2]_{FW} = 1.0$ ppm).

HWC, but becoming larger under NWC than would otherwise be the case in the absence of catalysis. This latter prediction is entirely consistent with the CGR versus ECP correlation shown in Fig. 16 and is of considerable importance, because, if a plant has to abandon HWC, for any reason, and the operators are unable to deactivate the catalyst on the steel surfaces, a period of enhanced CGR may ensue, resulting in a significant accumulation of damage.

In the case where the exchange current densities of the redox reactions are reduced by the presence of an inhibitive coating

(general inhibition, Figs 30 and 31), the ECP and CGR are predicted to decrease, regardless of whether excess hydrogen is present in the coolant. In the present calculations, with the assumed parameter values, the crack growth rate is predicted to be suppressed to the creep CGR value, except for the core channels. Thus, HWC

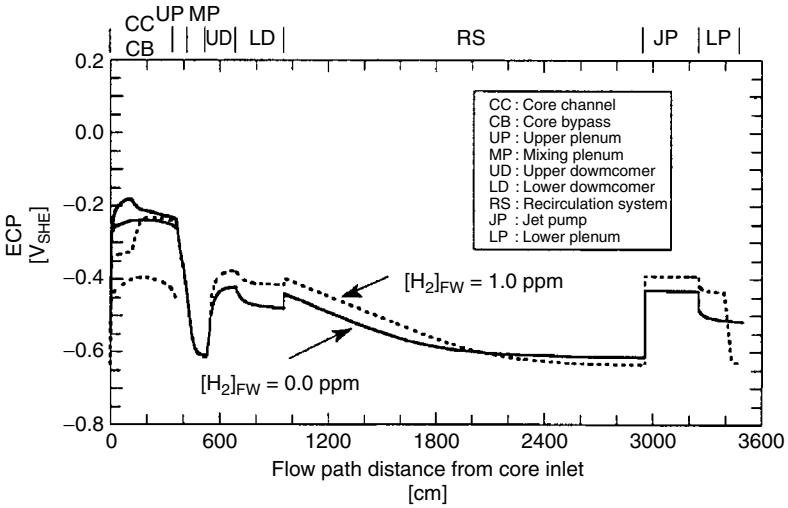


Fig. 30 ECP variation in the heat transport circuit of Dresden-2 employing general inhibition (exchange current density multiplier 10^{-4}) under NWC ($[H_2]_{FW}$ = 0.0 ppm) and HWC ($[H_2]_{FW}$ = 1.0 ppm).

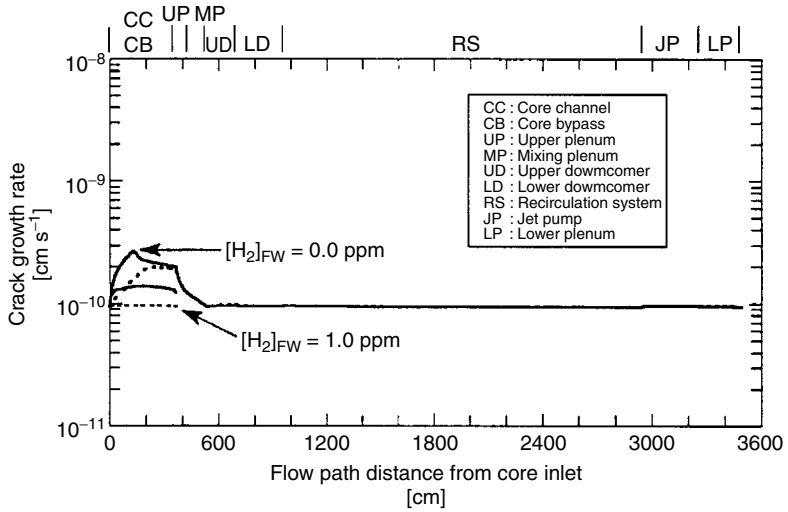


Fig. 31 Crack growth rate variation in the heat transport circuit of Dresden-2 employing general inhibition (exchange current density multiplier 10^{-4}) under NWC ($[H_2]_{FW}$ = 0.0 ppm) and HWC ($[H_2]_{FW}$ = 1.0 ppm).

becomes superfluous, and the operator might avoid the problems associated with the buildup of N^{16} and the cost of storing large amounts of hydrogen on-site. The challenge is to develop methods of coating the internal surfaces of BWR primary coolant circuits with a dielectric coating of sufficient resistance to reduce the interfacial impedance by a factor of at least 100.

Experimental work [53] has shown that electrophoretically deposited Zirconia (ZrO_2) coatings on sensitized stainless steel reduce the exchange current density for oxygen reduction by a factor of 100–1000, and that the measured impact of the coatings on the ECP and CGR (Fig. 32) in dilute sulfate solutions at 288 °C are in excellent agreement with the predictions of the MPM and CEFM, respectively. The general inhibition strategy would seem to be particularly attractive, because it does not require HWC and hence avoids the N^{16} problem and the cost of hydrogen storage and injection, but, to date, no effective method of applying a dielectric coating has been devised.

To the author's knowledge, the MPM and CEFM are the only models that could have predicted the effects of catalysis and inhibition, because they are the only models known to the author that explicitly consider the electrochemical kinetics of the redox reactions that occur on the steel surface. Again, we emphasize that the accumulation of damage due to SCC is primarily an electrochemical phenomenon, and any quantitative, deterministic theory must address the kinetics of the charge-transfer processes in the system.

Reference was made above to ULCO, since it is instructive to explore the veracity of this remedial measure. By way of background, IGSCC damage/conductivity correlations have been noted between damage accumulated in "good" plants that have operated with low coolant conductivity, and "bad" plants, in which the coolant conductivity was excessively high. The reader will recall from Sect. 9.4 that conductivity affects the CGR by allowing the coupling current exiting the crack to flow to a greater area on the external surface, where it is consumed by oxygen reduction. Thus, the external surface can support a larger coupling current, which translates into an enhanced CGR. The impact of coolant conductivity on accumulated damage for a crack adjacent to the H-3 weld on the inside surface of the core shroud of a GE Model 6 BWR under NWC over a 20-month period of constant, full power operation is shown

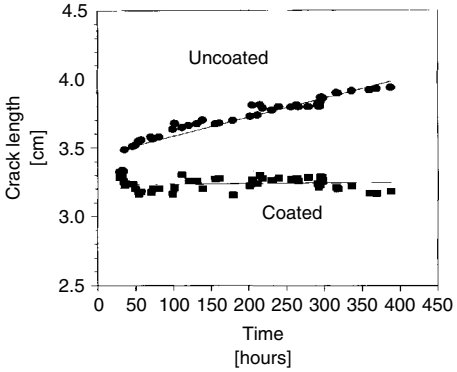


Fig. 32 Inhibition of IGSCC in sensitized Type 304 SS by a dielectric ZrO_2 coating. The experiments were carried out with two C(T) specimens daisy-chained together in the same solution (0.005 M Na_2SO_4) at 288 °C.

in Fig. 33. The conductivity values correspond to the ambient temperature (25 °C) values; of course, the values at the elevated temperature as calculated from the concentrations of ionic species and their limiting conductance values are employed in the calculation. The reader will note from Fig. 33 that coolant conductivity has a significant impact on the accumulated damage and that the crack length might be reduced from 1.9 cm to a little over 1.4 cm, from an initial value of 0.5 cm, by reducing the ambient temperature conductivity from 0.3 to 0.1 μS cm^{-1}. While significant, the gain is modest when compared against HWC, particularly with NMCA, and a detailed cost/benefit analysis would have to be carried out to determine whether the addition of the necessary facilities (additional ion exchange beds and larger flow fraction through the reactor water cleanup unit) are justified.

Another issue of great importance is the impact that stress has on the accumulation of IGSCC damage, because early efforts to mitigate cracking focused on relieving residual stresses in the HAZs adjacent to welds. For field welds, induction heating with water flowing through the pipe was used to render the stresses on the inner surface compressive. This treatment has an obvious benefit with regard to crack initiation (not treated in this chapter), but the question arises as to whether a comparable benefit exists for crack propagation. This question is important since, although a region below the pipe surface may be in compression, deeper regions are under tension and hence support IGSCC. Furthermore, under fatigue loading, even at high R ratio, a crack may grow through a compressive stress zone, resulting in a penetration of the pipe wall. The calculations presented

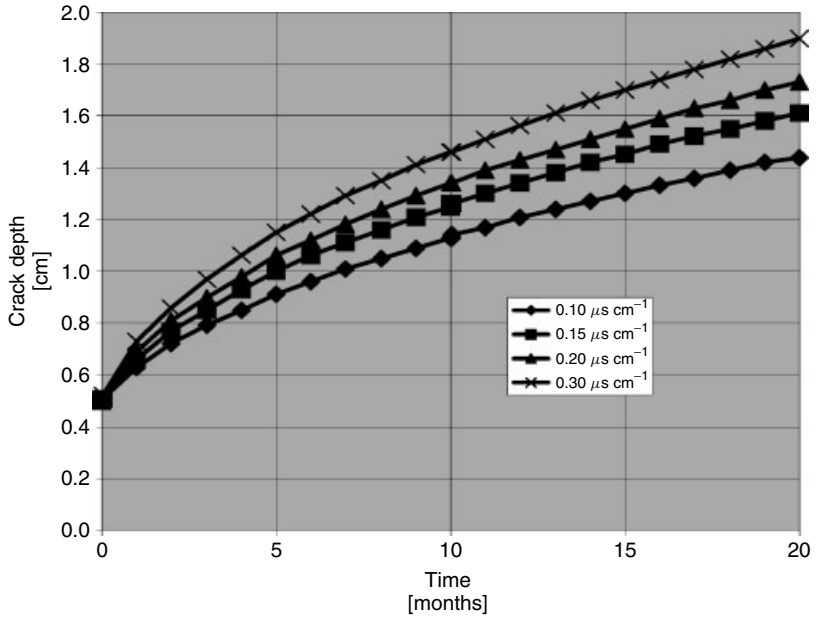


Fig. 33 Crack depth versus time curves for a crack at the H-3 weld in the shroud as a function of the coolant conductivity.

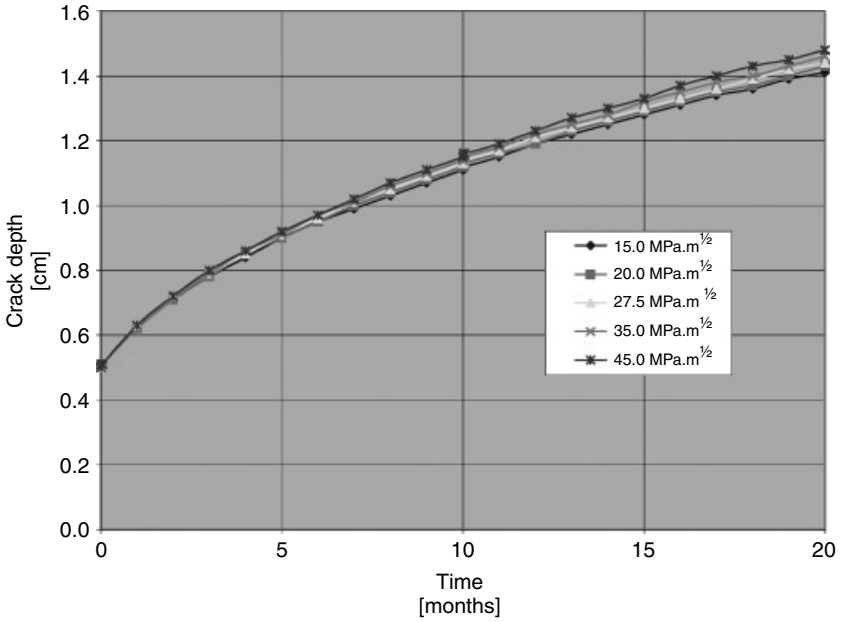


Fig. 34 Crack depth versus time curves for a crack at the H-3 weld in the shroud as a function of the stress intensity factor.

in Fig. 34 suggest that the accumulation of damage from a crack that has already penetrated the compressive zone is not a strong function of the stress intensity factor, corresponding to the fact that the crack exists within the Stage II region of the CGR versus K_I correlation. However, the author cautions that the initiation time probably dominates the total failure time, so that the lack of an effect of stress on the CGR should not be construed as indicating a lack of a dependence of the failure time upon stress. While the basic theory for the initiation of stress corrosion cracking has been developed in the form of damage function analysis [60–62], which incorporates the PDM [37] and the coupled environment pitting model (analogous to the CEFM, but for pit growth), and has been used to describe crack initiation in low pressure steam turbines, it has not been applied to crack initiation in primary coolant circuits of BWRs.

At this point, it is worth enquiring if these calculations have any relationship to reality. While it is very difficult to obtain information on crack extension in reactor coolant circuits for a variety of reasons, Tang et al. [63] published the data shown in Fig. 35. The data refer to the extension of a crack adjacent to the H-3 weld on the inner surface of the core shroud of a GE BWR in Taiwan. The authors had monitored the growth of the crack as a function of time after the eleventh outage for refueling. The reactor model was the same as that employed in our previous modeling and the coolant chemistry conditions could be estimated with sufficient accuracy to make a comparison between the observed and calculated crack extensions

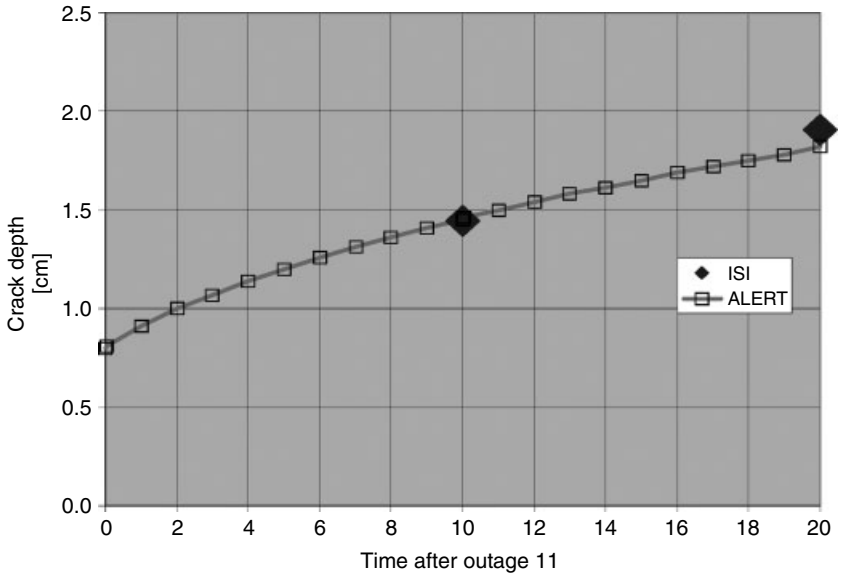


Fig. 35 Depth of a crack at the H-3 weld in the shroud as a function of time (months) after Outage 11. □ – values calculated by ALERT. ♦ – reported by Tang et al. [63].

meaningful. Note, however, that we did not know the exact number of outages, primarily unscheduled outages, but our previous work (Fig. 26) had shown that outages appear as small perturbations on the crack extension versus operating time correlation. Accordingly, any comparison would necessarily be approximate, but still meaningful. The remaining issue that needed to be resolved was the initiation time of the crack. This proved to be impossible to resolve, so the time of crack growth in the CGR versus time correlation calculated from ALERT was adjusted to ensure that the calculated and observed data coincided at 10 months after Outage 11. Thus, only the comparison at 20 months after Outage 11 has any probative value. As seen in Fig. 35, excellent agreement is obtained attesting to the accuracy of the code ALERT, given the uncertainty in the input data.

9.6 Electrochemistry of Pressurized Water Reactors

Compared with the work that has been reported on modeling BWR primary coolant circuits, relatively little work has been carried out to assess electrochemical effects in PWR primary circuits. This state of affairs reflects the fact that cracking has not been as great a problem in PWR primary coolant circuits as it has been in BWR primary coolant circuits. Secondly, because of the high hydrogen concentration (25 cc (STP)/kg (H_2O) −50 cc (STP)/kg (H_2O) corresponding to 1.12×10^{-3}–2.24×10^{-3} m) employed in a PWR primary circuit to ''suppress radiolysis,'' and in view of the lack of sustained boiling, it was generally believed that the ECP was dominated by the hydrogen equilibrium potential and hence that the coolant circuit acts as a ''giant

hydrogen electrode.'' If so, the potential is readily calculated from the pH of the coolant and the hydrogen concentration using the Nernst equation (Eq. 19). In the light of subsequent modeling, this picture is not entirely accurate, but more importantly PWRs are not free of cracking in their primary circuits and the cracking that is observed is very potential dependent. For example, primary water stress corrosion cracking (PWSCC) of Alloy 600 has plagued operators for many years and cracking of core barrel bolts (highly cold-worked Type 316 SS) has been a recurring problem. While there are significant materials differences between BWR and PWR primary circuits, in both cases it has gradually become evident that the electrochemistry of the coolant was a prime factor in the nucleation and propagation of cracks.

The primary coolant circuit of a PWR is shown in schematic form in Fig. 36. In this particular circuit, there are four loops between the reactor and the steam generators. The pressurizer is also shown, which maintains the pressure in the primary loop at a sufficiently high value (typically 150 bar) such that sustained boiling does not occur and maintains the desired concentration of hydrogen in the coolant. The reactor heat removal system (RHRS) and the reactor water cleanup system are not shown. The general operating conditions in a PWR primary loop are summarized in Table 2.

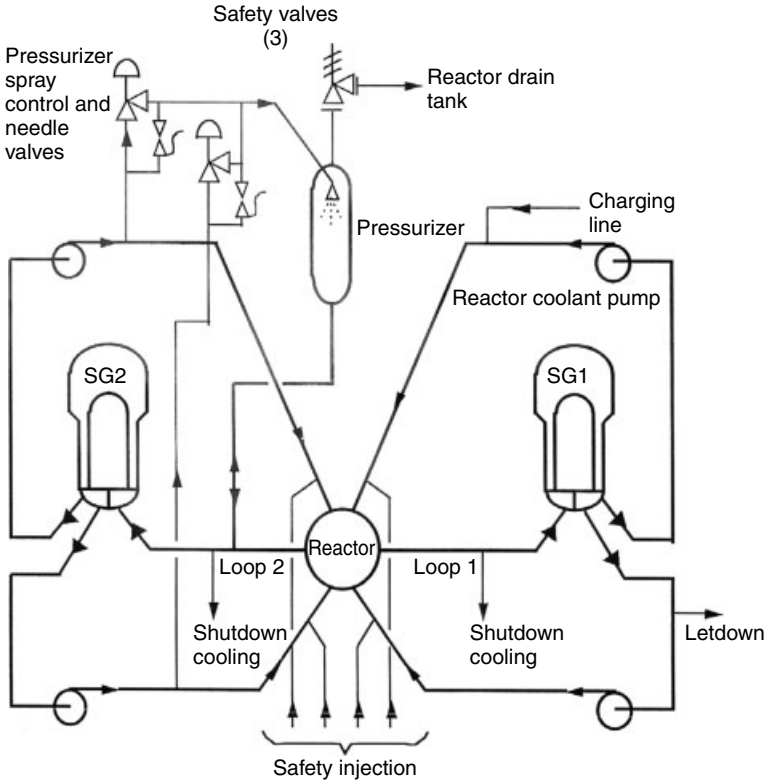


Fig. 36 Schematic of a PWR primary coolant circuit.

Tab. 2 Typical conditions that exist in the main loop of the primary coolant circuit of a PWR

Property	*Value*	*Comment*
Temperature	295–330 °C	Typical
Pressure	150 bar (2250 psi)	Typical
Coolant composition	4000–0 ppm B as boric acid, 4–1 ppm Li as lithium hydroxide, depending upon the burn-up of the fuel and the vendor.	Li–B trajectory over a typical fuel cycle is shown in Fig. 38
Hydrogen concentration	25–55 cc (STP)/kg (H_2O)	Some noncommercial units operate with [H_2] as high as 70 cc (STP)/kg (H_2O)
Core channel dose rate		Typical
γ photon	3×10^5 rad s^{-1}	
Neutron	6×10^5 rad s^{-1}	
α particles	3×10^5 rad s^{-1}	
Coolant mass flow rate	18 000 kg s^{-1}	Typical

The vessel of a PWR differs substantially from that of a BWR in that the control rods penetrate the vessel head, with only various instrumentation nozzles and assemblies penetrating the bottom (Fig. 37). Also, because the coolant does not boil (except for nucleate boiling in the fuel channels), the vessel does not contain steam separators or dryers, as in the case of a BWR. Accordingly, no separation of volatile gases to the steam phase occurs as it does in the case of a BWR. This makes describing the electrochemistry in a PWR primary coolant circuit somewhat more straightforward than in the case of a BWR.

There are significant differences between the primary coolant circuits of PWRs and BWRs, with regard to the coolant chemistry, physical geometry, flow path configuration, and in-core dose rates. First, the coolant in the primary circuit of a PWR is a boric acid/lithium hydroxide solution, with the boric acid concentration starting at the beginning of a fuel cycle at a concentration of 2000–4000 ppm of boron (0.18–0.36 M) and ending being essentially depleted. The boron is used as a ''nuclear shim'' to control core reactivity via the absorption of neutrons by $^{10}B_5$, which has a high capture cross section for thermal neutrons. Thus, as the $^{235}U_{92}$ and $^{239}Pu_{94}$ in the fuel undergoes fission (note that a significant fraction of the power at the end of the fuel cycle is produced by fission of $^{239}Pu_{94}$) and the concentration of $^{235}U_{92}$ in the fuel is depleted from an initial value of about 2.5% (compared with the natural abundance of 0.7%), the number of ''excess'' neutrons gradually decreases corresponding to a reduction in core reactivity, so the boron concentration is reduced correspondingly. On the other hand, the lithium concentration increase to 1–4 ppm over the same period, with the exact level being controlled by ion exchange, in order to control the pH and hence minimize corrosion, mass transport, and activity transport in the primary coolant circuit. A common trajectory for lithium versus boron over a fuel cycle is depicted by the hatched path in Fig. 38. It is also important to note that boron, which

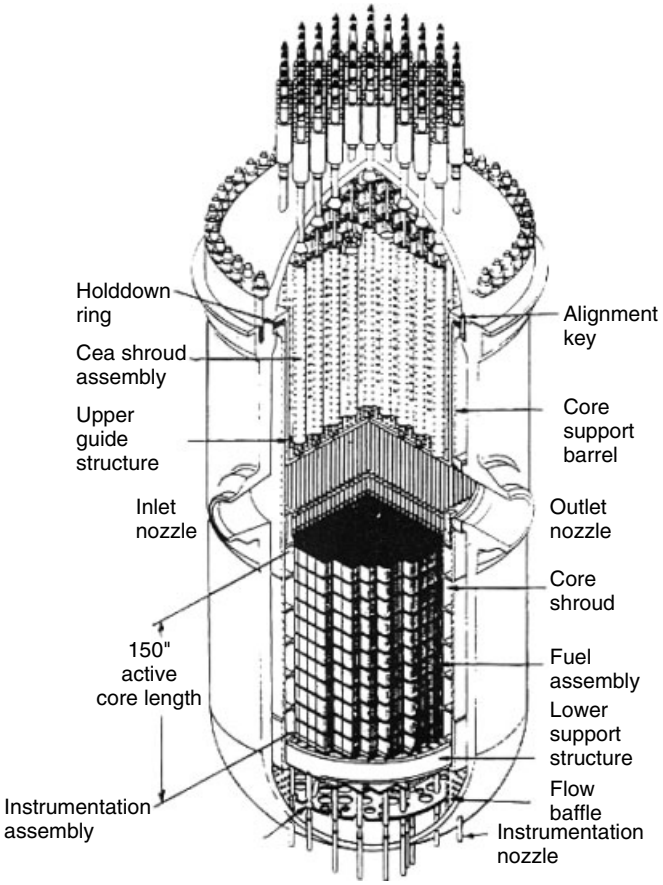


Fig. 37 Schematic of a PWR reactor vessel showing the various internal components.

is added to the primary coolant to act as a ''nuclear shim,'' in order to control nuclear reactivity by absorbing neutrons, as discussed above, produces lithium via the nuclear reaction $^{10}B_5(^1n_0, ^4He_2)^7Li_3$. The α particles ($^4He_2$) produced by this reaction contribute to the radiolysis of water (Table 1 summarizes the radiolytic yields of various radiolysis products for α-particle irradiation, compared with γ photons and neutrons). Indeed, typical dose rates of γ photons, neutrons, and α particles in PWR cores are 3×10^5 R s^{-1}, 6×105 rad s^{-1}, and 3×10^5 R s^{-1}, respectively (Table 2). These dose rates are about an order of magnitude greater than those in a BWR.

Modeling of the electrochemistry of a PWR primary coolant circuit has been achieved by Macdonald and Urquidi-Macdonald [18, 32, 34, 64] using a custom code titled PWR_ECP. This code is similar in structure to the DAMAGE-PREDICTOR code that we previously employed to model state points in BWRs, as described above. The code has not been developed to the point of predicting the accumulation of damage over the CEP for the reactor, primarily because of the lack of a viable CGR

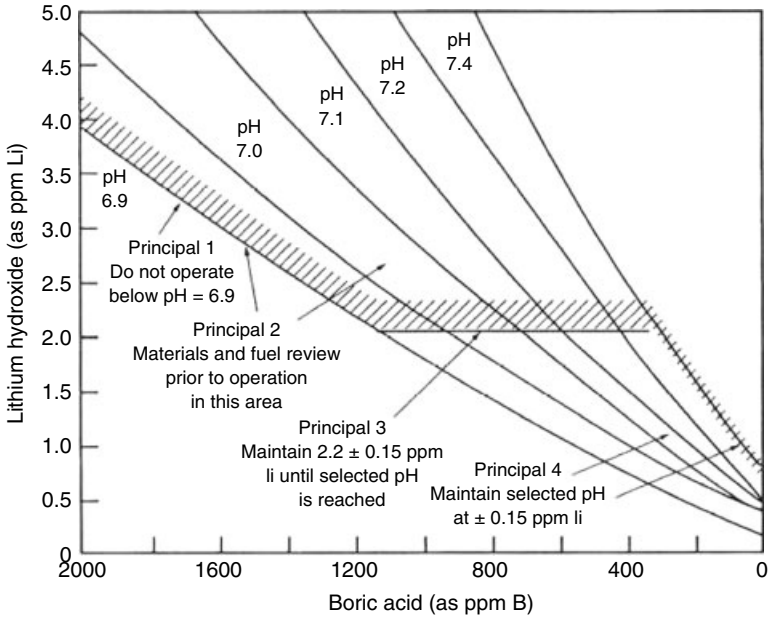


Fig. 38 pH control strategies affected by adjusting the lithium concentration as the boron is consumed during fuel burn-up. The trajectory that is commonly employed over a typical fuel cycle is represented by the hatched path, with the lithium concentration being controlled by ion exchange.

model that can accommodate hydrogen-induced cracking (HIC). (Such a model is now being developed in the authors' group and is expected to be introduced into PWR codes within the next two years). Instead, PWR_ECP predicts the ECP versus distance in the main loop and three secondary loops, as follows: (1) main loop, (2) RWCU loop, (3) RHRS, and (3) pressurizer.

The PWR_ECP code begins by employing a chemical speciation model to calculate the composition of the coolant and the pH at closely spaced points around the primary coolant circuit, including the pressurizer, RWCU, and RHRS loops (Fig. 39). Note that the purpose of the RHRS is to remove decay heat from the core upon shutdown of the reactor. Accordingly, this system is isolated from the primary coolant circuit during normal operation. The radiolysis of water is predicted using the full 48-model radiolysis model [32] with the radiolytic yields given in Table 1. The calculations were carried out at 10-cm increments, with the ECP being calculated at each increment using the MPM (Sect. 9.3). Dose rates for γ photons, neutrons, and α particles, coolant path geometry, and detailed thermal hydraulic data for each location were provided by the vendor for whom the modeling work was performed.

The primary coolant circuit loops were divided into 28 sections (cf, 10 or 11 sections in the BWR case) as shown in Fig. 39. Each section is characterized by an inlet and outlet temperature; γ, n, and α dose rates; flow fraction; linear flow velocity; hydrodynamic diameter; section length; material of construction (SS $\equiv$

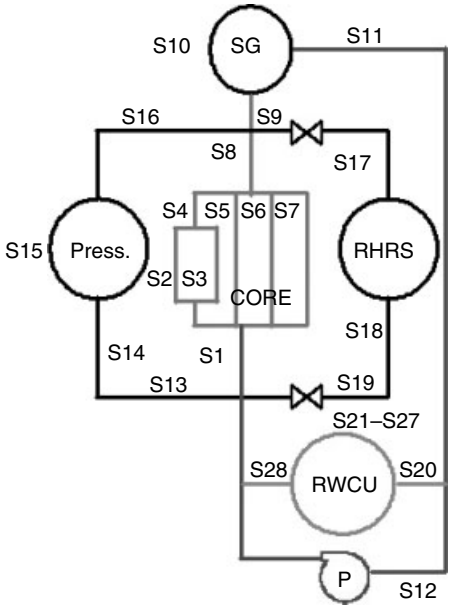


Fig. 39 Schematic of the primary coolant circuit of a PWR, showing the three principal loops: (1) the main loop (S1–S12), (2) the reactor heat removal system (RHRS, S17–S19), (3) the reactor water cleanup unit (RWCU, S20–S28), and (4) pressurizer (Press.: S13–S16).

stainless steel); and exchange current density multipliers, as in the case of the BWR models discussed earlier. If the inlet and outlet temperatures are different, it is assumed that the temperature varies linearly with distance between the two values. The code was used to model the ECP around the primary coolant system for normal operation with deoxygenated (5 ppb O_2) or oxygenated makeup water (5.0 ppm O_2) and for cold shutdown conditions. Facilities are included in the PWR_ECP code for modeling the periodic injection of oxygen and hydrogen peroxide into the RWCU and oxygen into the RHRS. The concentration of hydrogen was set to 55 cc (STP)/kg (H_2O) and the coolant mass flow rate was 17 880 kg s^{-1}.

Figure 40 displays the predicted ECP versus distance from the bottom of the core for full power, normal operation, which includes the injection of 5 mg kg^{-1} O_2 into the feedwater. All of the sections, except Section 7, display ECP values that lie between -0.35 V_{SHE} at the core entrance to -0.30 V_{SHE} at the core exit. On the other hand, Section 7, which is the bypass channel, is predicted to have an ECP of -0.75 V_{SHE}, regardless of the position. The higher ECP in the fuel channels and related areas reflects the generation of oxidizing radiolysis products, in spite of the fact that the coolant contains 55 cc (STP)/kg (H_2O).

Figures 41–43 display the predicted ECP versus distance for the three principal loops in the coolant circuit under normal, full power operation with oxygenated makeup water: (1) the main loop, (2) the pressurizer, and (3) the RWCU. Note that under these operating conditions, the RHRS is not operated, so calculated ECP data for this loop are not presented. The most important finding is that the ECP varies over a wide range (-0.74 to -0.20 V_{SHE} in the main loop) in a manner that is reminiscent of the primary coolant circuit of a BWR. Accordingly, the widespread belief that the primary circuit of a PWR behaves as a ''giant hydrogen electrode'' appears to be untenable, at least on the basis of the present work.

The ECP in the pressurizer circuit is predicted to be very negative over most of the loop, with values ranging from -0.75 to -0.81 V_{SHE} in Sections 13 and 16 Fig. 39. However, the ECP in Sections 14 and 15 Fig. 39 is predicted to rise to about -0.45 V_{SHE}. The origin of this rise is currently unknown.

Finally, the ECP in the RWCU is predicted to be very positive over most

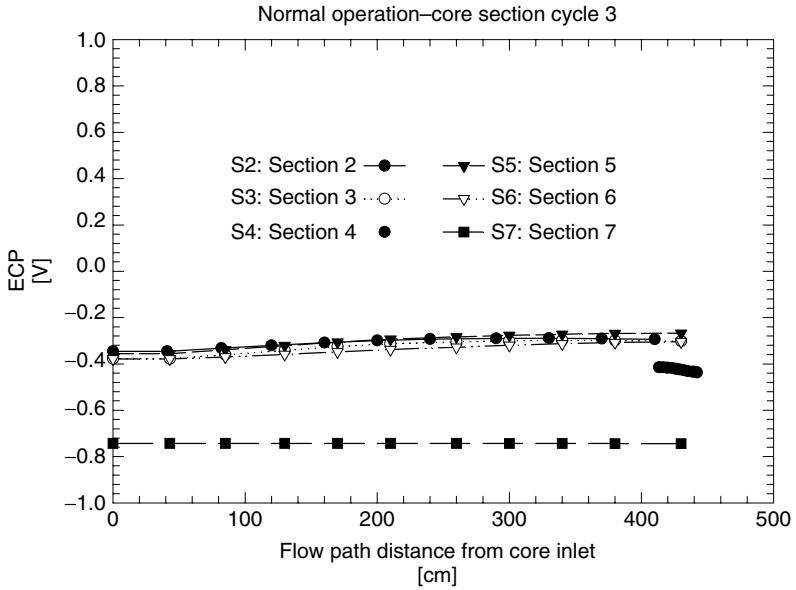


Fig. 40 ECP versus distance in a PWR under full Normal power operation. S1 = core return, S2 = fuel channels, S3 = hot zone, S4 = top part of the core, S5 = by-pass tube guide, S6 = by-pass hot zone, S7 = by-pass grid.

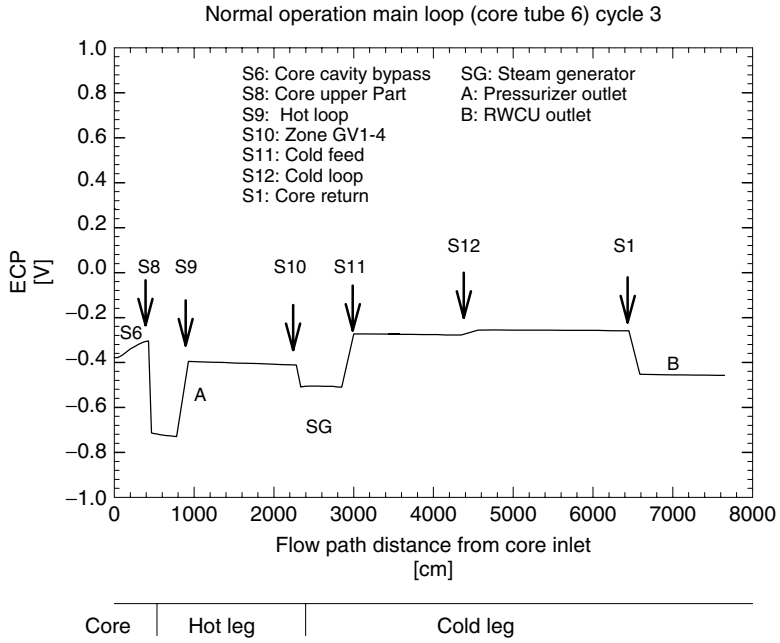


Fig. 41 ECP versus distance for the main primary coolant loop in a PWR under full Normal power operation.

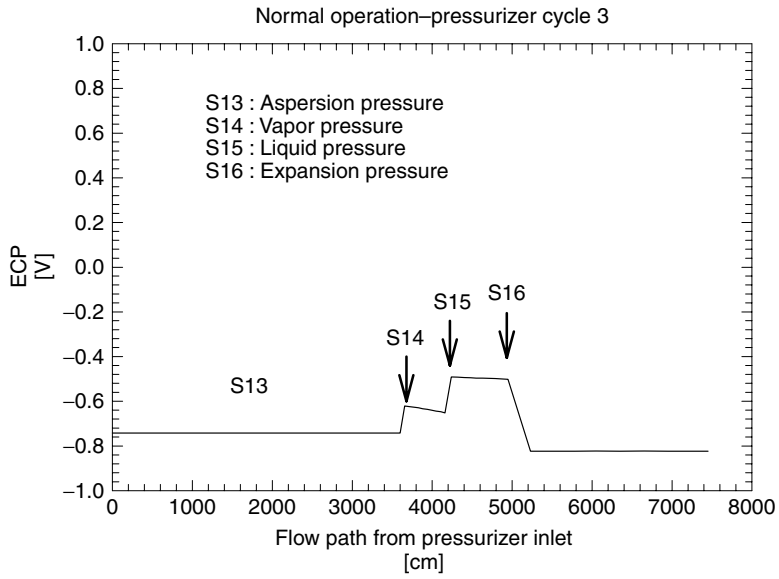


Fig. 42 ECP versus distance for the pressurizer loop in a PWR under full power operation.

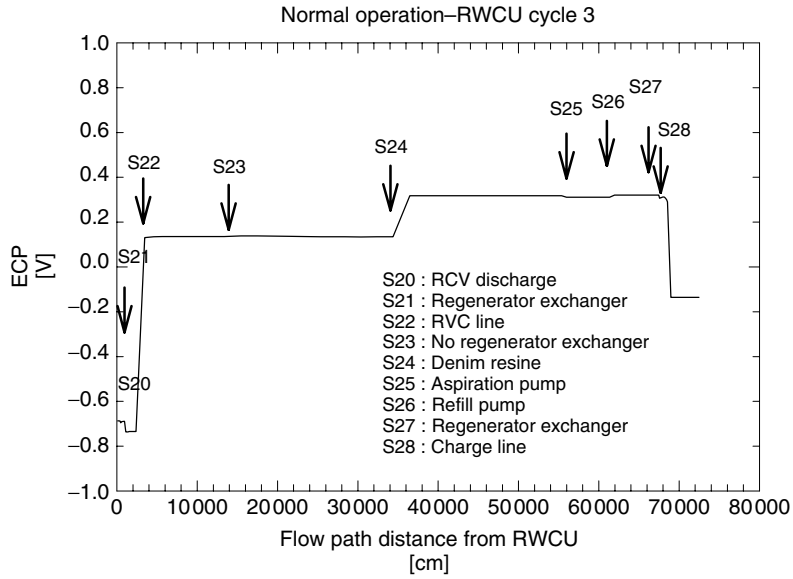


Fig. 43 ECP versus distance for the reactor water cleanup unit (RWCU) in a PWR under full power operation.

of the loop, corresponding to the low temperature of operation (45–55 °C). That the change in temperature is the origin of the sharp positive shift in the ECP is evident by comparing the ECP before and after the heat exchangers.

At this point, the reader might ask, ''Why is this modeling work important?'' It is important for at least two reasons: (1) much of the circuit is inaccessible to direct interrogation with regard to the ECP and hence the accumulation of damage and (2) many ''what if'' scenarios, which cannot be explored on operating reactors, because of the safety and economic issues involved, can be effectively investigated using sophisticated models of the type described in this chapter. Our work to date has emphasized BWRs, because that is where the greatest need has been in assisting plant operators to specify the most cost-effective operating protocols, but the work revealed here on PWRs demonstrates that electrochemical effects may determine the occurrence of certain damaging phenomena. These include the following:

1. Work carried out using this same model [16, 18, 32, 34] and the scoping calculations reported by Bertuch et al. [39] for PWR primary circuits indicates that, under certain circumstances, the ECP in the steam generators may become sufficiently negative that Alloy 600 (a common steam generator tubing material) could suffer HIC. Thus, the work of Totuska and Smialowska [65] indicates that cracking of Alloy 600 in PWR primary circuits occurs at potentials that are more negative than a critical value of −800 mV. This critical potential is perilously close to the calculated ECP in the steam generators, particularly when using deoxygenated makeup water. The calculations of Bertuch et al. [39] suggest that the ECP may become as negative as $-810\ mV_{SHE}$, which would pose a threat to steam generator integrity. Accordingly, we have argued for examination of a ''low hydrogen'' water chemistry, which would maintain the ECP at a value that is more positive than the critical value for hydrogen-induced fracture in Alloy 600 under all operational circumstances. A hydrogen concentration of about 5 cc (STP)/kg (H_2O) is believed to be sufficiently low to avoid primary side cracking of steam generator tubes [39].
2. Mass transport and the closely related activity transport phenomenon are fundamentally electrochemical in nature. Thus, to a significant extent, the movement of corrosion products around the primary circuits of both BWRs and PWRs is due to differential solubility of magnetite (Fe_3O_4), for example, induced by the variation in the ECP, coupled with hydrodynamic factors, as shown by the work of Salter-Williams [66] and Urquidi-Macdonald, Pitt, and Macdonald [32], for example. In the opinion of the author, mass and activity transport phenomena in WCNRs can only be understood and predicted by first defining the electrochemistry of the primary HTCs.

The ultimate goal would be to specify a corrosion evolutionary path of the type shown in Fig. 38 that would minimize the accumulated damage, which may be of different forms for different systems and operating protocols.

9.7 Summary

The application of deterministic models for simulating SCC phenomena in

BWR primary coolant circuits is described. The first-generation code, DAMAGE-PREDICTOR, has been used to model the radiolysis of the coolant, to estimate the ECP, and to calculate the CGR at fixed state points during reactor operation in 14 plants worldwide. This code has been validated in ''double-blind'' comparisons between the calculated and measured hydrogen concentration, oxygen concentration, and ECP in the recirculation system of the Leibstadt BWR in Switzerland, as well as through less formal comparisons with data from other plants. Second-generation codes have now been developed, including REMAIN for simulating BWRs with internal coolant pumps and the ALERT series for modeling reactors with external pumps, and a third-generation code, FOCUS, has just been completed. Two of this series, ALERT and FOCUS, yield the IDF, which is the crack length, on a component-by-component basis for a specified future operating scenario. This code therefore allows one to explore proposed future operating protocols, with the objective of identifying those that are most cost-effective and which minimize the risk of failure of components by SCC. The application of these codes is illustrated by exploring the benefits of partial HWC for an actual reactor, in which hydrogen is added to the feedwater over only limited periods during operation. The simulations show that the benefits, in terms of reduction in the IDFs for various components, is sensitive to when HWC was initiated in the plant life and to the length of time over which it is applied.

This modeling technology has now been extended to PWRs with the development of the code PWR_ECP. This code models the primary cooling circuit of the reactor comprising the main coolant loop (MCL), the pressurizer, the RWCU, and the RHRS. The latter does not operate while the reactor is under power. PWR_ECP calculates the corrosion potential at closely spaced points around the four loops for any given set of operating conditions. The principal finding is that the internal surfaces of a PWR primary coolant circuit do not act as a ''giant hydrogen electrode'' as is commonly assumed and, even at high hydrogen concentrations (55 cc (STP)/kg (H_2O)) surprisingly large variations in the ECP exist around the circuits that can be accounted for in terms of the variation in temperature and the generation of electroactive species via the radiolysis of water. The variations in the ECP and temperature are postulated to provide the driving force for operational problems, such as mass and activity transport and SCC (HIC) in the primary coolant circuits of PWRs.

References

1. P. Cohen, *Water Coolant Technology of Power Reactors*, Gordon and Breach Science Publishers, New York, 1969.
2. C.-C. Lin, *Radiochemistry in Nuclear Power Reactors*, NAS-NS-3119, National Academies Press, Washington, 1996.
3. M. E. Indig, J. L. Nelson, *Corrosion* **1991**, *47*, 202.
4. H. Cristensen, *Nucl. Technol.* **1995**, *109*, 373.
5. W. G. Burns, J. Henshaw, J. A. B. Goodball. *The Radiation Chemistry of a Pressurized Water Reactor. The Effect of Added Hydrogen Based on a Simplified Modeling Approach*, AEA Technology. Report AEA RS 3478, 1994.
6. T. K. Yeh, D. D. Macdonald, A. T. Motta, *Nucl. Sci. Eng.* **1995**, *121*, 468.
7. M. L. Lukashenko, A. J. Milaer, *Atomn. Energ.* **1992**, *72*, 570.
8. W. G. Burns, P. B. Moore, *Radiat. Eff.* **1976**, *30*, 233.
9. D. D. Macdonald, A. Bertuch, J. Pang et al., Estimation of corrosion potentials in the heat transport circuits of LWRs, *Proceedings, International Conference in Chemistry in Water Reactors: Operating Experience and*

New Developments, Nice, France, Apr. 24–27, 1994.
10. C. C. Lin, H. R. Helmholtz, *Int. J. Chem. Kinet.* **1991**, *23*, 971.
11. E. Ibe, M. Sakagami, S. Uchida, *Nucl. Sci. Technol.* **1986**, *23*, 11.
12. J. Chun, *Modeling of BWR Water Chemistry*, Master Thesis, Department of Nuclear Engineering, Massachusetts Institute of Technology, 1990.
13. T. K. Yeh, D. D. Macdonald, A. T. Motta, *Nucl. Sci. Eng.* **1996**, *123*, 295.
14. T. K. Yeh, D. D. Macdonald, A. T. Motta, *Nucl. Sci. Eng.* **1996**, *123*, 305.
15. C. P. Ruiz, C. C. Irvin, *Modeling Hydrogen Water Chemistry for BWR Applications*, EPRI-NP-6386, Electric Power Research Institute, June 1989.
16. K. Ishigure, *Rad. Phys. Chem.* **1987**, *29*, 195.
17. E. Ibe, *J. Nucl. Sci. Technol.* **1987**, *24*, 220.
18. D. D. Macdonald, M. Urquidi-Macdonald, J. H. Mahaffy et al., *Electrochemistry of Water-Cooled Nuclear Reactors*, Nuclear Energy Education Research, Department of Energy, Washington, D.C., Grant No. DE-FG07-021D14334, Final Report 2006.
19. D. D. Macdonald, I. Balachov, G. Engelhardt, *Power Plant Chemistry* **1999**, *1*, 9.
20. T.-K. Yeh, C.-T. Liu, C.-H. Tsai, *J. Nucl. Sci. Technol.* **2005**, *42*, 809.
21. T.-K. Yeh, F. Chu., *Nucl. Sci. Eng.* **2001**, *139*, 221.
22. T.-K. Yeh, C.-H. Tsai, Y.-H. Cheng, *J. Nucl. Sci. Technol.* **2005**, *42*, 462.
23. T.-K. Yeh, F. Chu, *J. Nucl. Sci. Technol.* **2000**, *37*, 1063.
24. T.-K. Yeh, *Nucl. Sci. Eng.* **2002**, *142*, 220.
25. T.-K. Yeh, C.-H. Tsai, *J. Nucl. Sci. Technol.* **2004**, *41*, 116.
26. T.-K. Yeh, C. Chang, F. Chu, C.-S. Huang, The predicted effectiveness of noble metal treatment at the chinshan boiling water reactor, *Proceedings of 9th International Symp. Env. Deg. Mat. Nuc. Pwr. Sys.-Water Reactors*, New Portbeach, TMS, ANS, NACE Int., (August 1999), p 1211.
27. T.-K. Yeh, M.-S. Yu, D. D. Macdonald, The effect of catalytic coatings on IGSCC mitigation for boiling water reactors operated under hydrogen water chemistry, *Proceedings of the 8th International Symposium Env. Deg. Mat. Nuc. Pwr. Sys.-Water Reactors*, Amelia Island (NACE Int.) August 1997.
28. T.-K. Yeh, K.-T. Lee, C.-H. Tsai, *Mech. Mat. Eng. Sci. Expts* **2003**, *41*, 461.
29. T.-K. Yeh, C.-H. Liang, M.-S. Yu, D. D. Macdonald, The effect of catalytic coatings on IGSCC mitigation for boiling water reactors operated under hydrogen water chemistry. *Proceedings 8th Int'l. Symp. Env. Deg. of Mat. Nuc. Pwr. Sys.–Water Reactors*. August 1995. Amelia Island (NACE International) 1997.
30. D. Macdonald, M. Urquidi-Macdonald, *Corrosion* **1996**, *52*, 659–670.
31. J. Elliot, *Rate Constants and G-Values for the Simulation of the Radiolysis of Light Water Over the Range 0–300°C*, Atomic Energy of Canada Ltd (AECL) Report No. 11073 October 1994.
32. D. D. Macdonald, J. Pitt, M. Urquidi-Macdonald, *Electrochemical Modeling of Activity Transport in Pressurized Water Reactors*, ICONE-13, Bejing, China, Paper 50423, May 16–20, 2005.
33. K. Makela, T. Buddas, M. Zmitko, J. Kysela, The effect of hydrazine on high temperature water chemistry and corrosion product transport in primary circuit of a VVER 440 unit, *Water Chemistry Conference*, Nice, France, 1994.
34. M. Urquidi-Macdonald, J. Pitt, D. D. Macdonald, Fundamental Aspects of Water Coolant, SKJ Report, (Ed.: Hilbert Christensen) ISSN 1104–1374, April 2006.
35. D. D. Macdonald, *Corrosion* **1992**, *48*, 194.
36. J. B. Lee, *Electrochemical Approach to Corrosion Problems of Several Iron-Nickel-Chromium alloys in High Temperature, High Pressure Water*, Ph D Dissertation, Ohio State University, Columbus, OH, 1978.
37. D. D. Macdonald, *Pure Appl. Chem.* **1999**, *71*, 951.
38. D. D. Macdonald, A. Sun, N. Priyantha, P. Jayaweera, *J. Electroanal. Chem.* **2004**, *572*, 421.
39. A. Bertuch, J. Pang, D. D. Macdonald, The argument for low hydrogen and lithium operation in PWR primary circuits. *Proceedings of 7th Int'l. Symp. Env. Deg. Mat. Nuc. Pwr. Sys.–Water Reactors*, Breckenridge, (NACE, Int.) 2, 687–697 (August 1995).
40. P. L. Andresen, F. P. Ford, *Mat. Sci. Eng., A* **1988**, *103*, 167.
41. D. D. Macdonald, M. Urquidi-Macdonald, *Corros. Sci.* **1991**, *32*, 51.
42. D. D. Macdonald, M. Urquidi-Macdonald, An advanced coupled environment fracture

model for predicting crack growth rates, *Proc. TMS Parkins Symp. Fund. Aspects Stress Corros. Crack*, Cincinnati, Oct. 20–24, 1991, pp. 443–455; TMS, Warrendale, 1992.
43. D. D. Macdonald, P. C. Lu, M. Urquidi-Macdonald et al., *Corrosion* **1996**, *52*, 768.
44. D. D. Macdonald, *Corros. Sci.* **1996**, *38*, 1033.
45. M. Vankeerbergen, D. D. Macdonald, *Corros. Sci.* **2002**, *44*, 1425.
46. G. R. Engelhardt, M. Urquidi-Macdonald, D. D. Macdonald, *Corros. Sci.* **1999**, *41*, 2267.
47. (a) S. Gavrilov, M. Vankeerbergen, G. Nelissen et al., Finite element calculation of crack propagation in type 304 stainless steel in dilute sulphuric acid solutions, *Corros. Sci.* **2006**, in press; (b) S. Gavrilov, *Finite Element Simulation of Stress Corrosion Cracking in an Aqueous Environment*, Ph.D. Dissertation, Vrije Universisteit Brussels, Belgium, 2003.
48. Q. J. Peng, J. Kwon, T. Shoji, *J. Nucl. Mater.* **2004**, *324*, 52.
49. P. Andresen, *Corrosion* **1993**, *49*, 714.
50. A. Turnbull, *Corros. Sci.* **1997**, *39*, 789.
51. D. S. Wilkinson, V. Vitek, *Acta Metall.* **1998**, *30*, 723.
52. M. P. Manahan Sr, D. D. Macdonald, A. J. Peterson Jr, *Corros. Sci.* **1995**, *37*, 189.
53. X. Zhou, I. Balachov, D. D. Macdonald, *Corros. Sci.* **1998**, *40*, 1349.
54. D. D. Macdonald, *The Electrochemistry of IGSCC Mitigation in BWR Coolant Circuits*, Chimie 2002, Water Chemistry in Nuclear Reactor Systems, Avignon, France (April 22–26, 2002).
55. D. D. Macdonald, G. Cragnolino, *The Development of Deterministic Methods for Predicting Corrosion Damage in Water Cooled Nuclear Reactors*, PSU, Final Report to ESEERCO, EP 93-33, (November, 1996).
56. D. D. Macdonald, G. Cragnolino. The Critical Potential for the IGSCC of Sensitized Type 304 SS in High Temperature Aqueous Systems. *Proceedings 2nd Int'l. Symp. Env. Deg. Mat. Nucl. Power Syst.–Water Reactors*, Monterey, ANS. (September 9–12, 1985).
57. J. Congleton, T. Shoji, R. N. Parkins, *Corros. Sci.* **1985**, *25*, 633.
58. D. D. Macdonald, L. Kriksunov, Flow rate dependence of localized corrosion processes in thermal power plants, *Advances in Electrochemical Science and Engineering* Vol. 5, John Wiley and Sons, New York, 1997, 125–193.
59. Y.-J. Kim, L. W. Niedrach, P. L. Andresen, *Corrosion* **1996**, *52*, 738.
60. G. Englehardt, M. Urquidi-Macdonald, D. D. Macdonald, *Corros. Sci.* **1997**, *39*, 419.
61. D. D. Macdonald, M. Urquidi-Macdonald, *Corrosion* **1992**, *48*, 354.
62. D. D. Macdonald, C. Liu, M. Urquidi-Macdonald et al., *Corrosion* **1994**, *50*, 761.
63. [65] J. R. Tang, L. Kao, D. Y. Shiau et al., *Nucl. Technol.* **1998**, *121*, 324.
64. M. Urquidi-Macdonald, D. D. Macdonald, *The PWR_ECP Code for Modeling the Impact of the Feedwater Oxygen on the Primary Circuit ECP*, Internal Report, Framatome, (June 2000).
65. N. Totsuka, Z. Szklarska-Smialowska, *Corrosion* **1987**, *43*, 734.
66. M. Salter-Williams, *Mass Transport of Magnetite in Pressurized Water Reactor Primary Circuits*, M.S. Thesis, Department Nuclear Eng., Penn State University, 2000.

10 Electrochemical Techniques for Studying High-temperature Subcritical and Supercritical Aqueous Solutions

Serguei N. Lvov
The Pennsylvania State University University Park, PA, USA

10.1 Introduction **725**

10.2 High-temperature Electrochemical Systems **726**
10.2.1 High-temperature Reference Electrodes 730
10.2.2 High-temperature Indicator Electrodes 732
10.2.3 Undesirable Potentials 735
10.2.4 High-temperature Reference and Buffer Solutions 736

10.3 High-temperature Potentiometry **736**

10.4 High-temperature pH Measurements **739**

10.5 High-temperature Electrokinetic Studies **741**

10.6 High-temperature Electrochemical Kinetics and Corrosion **741**

10.7 High-temperature Electrical Conductance Techniques **744**

10.8 Conclusions **745**
References 746

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

10.1 Introduction

Water is the most ubiquitous of solvents, and by virtue of its extraordinary physicochemical and transport properties, it forms the medium in which processes as diverse as biochemical and geochemical occur. The particular features of high-temperature aqueous solutions have been widely recognized [1] and employed in many areas of science and technology, such as thermal power generation, supercritical water oxidation of organic waste, hydrothermal synthesis of ceramic materials, fluid sedimentary basins, nuclear waste disposal, and so on. However, owing to the difficulty of performing experimental measurements at high temperature and pressure, there is a scarcity of electrochemical studies under high-temperature subcritical and supercritical conditions (critical point of water is found at 374 °C and 221 bar).

High-temperature aqueous solutions represent new frontiers in electrochemical studies that are both technically challenging and technologically important [2]. The interest in high-temperature aqueous electrochemistry has significantly increased in the last 30 years and this interest has risen mainly because of the many technological processes that take place in the high-temperature aqueous environment. However, it is still difficult to routinely carry out high-temperature electrochemical studies because only a few well-established electrochemical techniques are available for performing high-temperature electrochemical measurements. Additionally, the lack of an inert isolating material that can be used at temperatures above the upper stability limit of Teflon (300 °C) severely hinders electrochemists who deal with high-temperature subcritical and supercritical electrochemical systems.

This article details the thus far developed experimental techniques to carry out potentiometric, pH, electrokinetic, electrochemical kinetics, corrosion, and conductivity measurements in high-temperature (>300 °C) subcritical and supercritical aqueous environments. The author of this chapter recently reviewed the electrochemical processes in high-temperature aqueous solutions [2], an experience that has had a significant impact on the content of this chapter. Note that the treatment and interpretation of the obtained high-temperature electrochemical data are out of the scope of this review, but there are a number of excellent papers [3–6], which the author recommends to a reader who is interested in the treatment of electrochemical data. Also, two of these papers [4, 5] are useful to anyone interested

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

in electrochemical techniques to be used at elevated temperatures below 300 °C.

10.2 High-temperature Electrochemical Systems

In general, electrochemical systems are heterogeneous and involve at least one (or both) of the fundamental processes – mass transport and an electron-transfer reaction. Moreover, electrochemical reactions involve charged species, so the rate of the electron-transfer reaction depends on the electric potential difference between the phases (e.g. between the electrode surface and the solution). The mass transport processes mainly include diffusion, conduction, and convection, and should be taken into account if the electron-transfer reaction properties are to be extracted from the experimental measurements. The proper control of the mass transport processes seems to be one of the main problems of high-temperature electrochemical studies.

The main property of the ith component in an electrochemical system is the electrochemical potential, $\tilde{\mu}_i^{\alpha}$, which is defined as

$$\tilde{\mu}_i^{\alpha} = \mu_i^{\alpha} + z_i F E^{\alpha} \tag{1}$$

where μ_i^{α} and z_i are the chemical potential and the charge of species, i, respectively; F is the Faraday constant (96 485 C mol^{-1}); and E^{α} is the inner electric potential in phase α. Although E^{α} is not a measurable quantity, the difference, E, between the inner electric potentials of two electrically conducting phases can be either measured or established in electrochemical experiments. In addition, a part of the inner electric potential, which is called the *zeta potential*, can be obtained experimentally from electrokinetic studies of a solid/liquid interface. Zeta potential is one of the key parameters of the electrical double layer (EDL) (see Sect. 10.5), and can be used in modeling the EDL structure.

A particular feature of the electrochemical technique is that the total rate of the electrochemical process can be defined by measuring the current density, j, flowing in the electrical circuit where the reaction rate, v (per unit of surface area), is related to the current density as follows:

$$j = nFv \tag{2}$$

where n is the number of electrons involved in the electrochemical reaction.

By considering a reduction/oxidation electrode reaction as a first-order interfacial process [7]:

$$O + ne^- \underset{k_b}{\overset{k_f}{\longleftrightarrow}} R \tag{3}$$

the reaction rate will be defined (1) by the heterogeneous rate constants for the reduction, k_f, and oxidation, k_b, processes; and (2) by the activities of O, ${a_O}^{\dagger}$, and R, ${a_R}^{\dagger}$, at the electrode surface as follows:

$$v = k_f \, {a_O}^{\dagger} - k_b \, {a_R}^{\dagger} \tag{4}$$

Within the framework of the activation complex theory, the rate constants can be related to the Gibbs energies of activation for the reduction, $\Delta {G_R}^{\dagger}$, and oxidation, $\Delta {G_O}^{\dagger}$, reactions as follows:

$$k_f = A_f \exp\left(\frac{-\Delta {G_R}^{\dagger}}{RT}\right) \tag{5}$$

$$k_b = A_b \exp\left(\frac{-\Delta {G_O}^{\dagger}}{RT}\right) \tag{6}$$

where A_f and A_b are the preexponential factors, R is the gas constant (8.3145 J·K^{-1}·mol^{-1}) and T is the temperature in Kelvin.

The Gibbs energies of activation for the reduction and oxidation processes are different and depend on the potential difference, E, and the potential barrier symmetry, which is defined by the transfer coefficients α_{red} and α_{oxid} for the reduction and oxidation reactions respectively. For a symmetrical barrier, $\alpha_{red} = \alpha_{oxid} = 0.5$, so for a simple reaction, $\alpha_{red} + \alpha_{oxid} = 1$, but generally the transfer coefficients are between 0 and 1. The relationships between the Gibbs energies of activation, the transfer coefficients, and the potential difference between phases are well known and can be presented for the reduction and oxidation processes as follows:

$$\Delta G_R^{\dagger} = \Delta G_R^{o\dagger} + \alpha_{red} nFE \qquad (7)$$

$$\Delta G_O^{\dagger} = \Delta G_O^{o\dagger} - \alpha_{oxid} nFE \qquad (8)$$

where $\Delta G_R^{o\dagger}$ and $\Delta G_O^{o\dagger}$ are the Gibbs energies of activation for the reduction and oxidation processes, respectively, if the potential difference between the phases is zero. The values of $\Delta G_R^{o\dagger}$ and $\Delta G_O^{o\dagger}$ can also be defined by the corresponding rate constants, k_f^o and k_b^o, as follows:

$$k_f^o = A_f \exp\left(\frac{-\Delta G_R^{o\dagger}}{RT}\right) = k_f \exp\left(\frac{\alpha_{red} nFE}{RT}\right) \qquad (9)$$

$$k_b^o = A_b \exp\left(\frac{-\Delta G_O^{o\dagger}}{RT}\right) = k_b \exp\left(\frac{-\alpha_{oxid} nFE}{RT}\right) \qquad (10)$$

If the standard state conditions prevail, that is, $E = E^{\circ}$, then, $k_f^o \exp[-\alpha_{red} nF(E - E^{\circ})/RT] = k_b^o \exp[\alpha_{oxid} nF(E - E^{\circ})/RT] = k^o$, where k^o is a key characteristic of the electron-transfer reaction and is called the *standard rate constant*. The standard rate constant can be used to relate the rate constants, k_f and k_b, to the potential difference, E, and its equilibrium value, E°, as follows:

$$k_f = k^o \exp\left[\frac{-\alpha_{red} nF(E - E^{\circ})}{RT}\right] \qquad (11)$$

$$k_b = k^o \exp\left[\frac{\alpha_{oxid} nF(E - E^{\circ})}{RT}\right] \qquad (12)$$

Insertion of these relationships into Eqs (2) and (4) yields the complete current–potential characteristic, j, [7]:

$$j = nFk^o \left\{ a_O{}^{\dagger} \exp\left[\frac{-\alpha_{red} nF(E - E^{\circ})}{RT}\right] - a_R{}^{\dagger} \exp\left[\frac{\alpha_{oxid} nF(E - E^{\circ})}{RT}\right] \right\} \qquad (13)$$

This is a general equation that should be used in the development of high-temperature electrochemical kinetics and corrosion measurements if the processes of mass transport are already taken into account.

The physical meaning of k^o relates to the kinetic ability of a reducing/oxidizing electrochemical system. In other words, a system with a large k^o will reach equilibrium quickly, whereas a system with a small k^o is sluggish. For carrying out high-temperature potentiometric studies ($j = 0$) and implementing new high-temperature electrodes, one should find the k^o for the electrochemical reaction of interest to be sure that the chosen electrode is suitable for such potentiometric measurements. However, there is a significant deficit of such information for high-temperature electrochemical systems.

At equilibrium, $j = 0$, and the potential difference will approach the equilibrium value, E^{eq}, so that Eq. (13) can be written

as

$$nFk^{\circ}{a_{\mathrm{O}}}^{\dagger}\exp\left[\frac{-\alpha_{\mathrm{red}}nF(E^{\mathrm{eq}}-E^{\circ})}{RT}\right] = nFk^{\circ}{a_{\mathrm{R}}}^{\dagger}\exp\left[\frac{\alpha_{\mathrm{oxid}}nF(E^{\mathrm{eq}}-E^{\circ})}{RT}\right] = j_{\mathrm{o}} \quad (14)$$

The value j_{o} is the exchange current density, which is an important characteristic of the electrochemical system and its reversibility. Taking into account that at equilibrium the bulk activities of O, a_{O}, and R, a_{R}, are the same as at the surface, Eq. (14) can be converted to the well-known Nernst equation:

$$E^{\mathrm{eq}} = E^{\circ} + \left(\frac{RT}{nF}\right)\ln\left(\frac{a_{\mathrm{O}}}{a_{\mathrm{R}}}\right) \quad (15)$$

for the electrochemical reaction presented by Eq. (3). This equation forms the basis for potentiometric measurements so that electrode systems are often referred to as exhibiting Nernstian behavior or response during their calibration when the measured potential, E^{eq}, is a linear function of $\ln(a_{\mathrm{O}}/a_{\mathrm{R}})$ with a ''theoretical'' slope of RT/nF as dictated by Eq. (15).

Taking thermodynamics into account, the electric potential of an individual electrode, E_i^{eq}, can be related to the Gibbs energy, $\Delta_r G_i$, of the corresponding electrochemical half-reaction as follows:

$$\Delta_r G_i = -z_i F E_i^{\mathrm{eq}} \quad (16)$$

where z_i is the charge number of a half-cell electrochemical reaction. Conventionally, the standard values of both $\Delta_r G_i$ and E_i^{eq} are zero at all temperatures ($\Delta_r G_i^0 = 0$, $E_i^0 = 0$) for the hydrogen electrode reaction (HER), $2H^+(aq) + 2e^- \leftrightarrow H_2(g)$, so that all other electrode potentials can be estimated with respect to the standard hydrogen electrode (SHE) scale. Note that by convention the Gibbs energy of formation of the electron is also assigned to be zero at all conditions so that the thermodynamic properties of the electron are never shown in connection with the electrochemical half-reactions.

A significant quantity of the standard electrode potentials, E_i^0, has been experimentally estimated at ambient conditions (25 °C and 1 bar) [8]; however, only a limited number of E_i^0 values can be found at high temperatures and pressures [9]. Therefore, experimental studies to derive a comprehensive list of E_i^0 values for a wide range of temperatures are a serious challenge in high-temperature experimental electrochemistry.

In general, the open-circuit potential measured between two reversible electrodes, which is also called *electromotive force*, E^{eq}, is defined by the Nernst equation. A simplified form of this equation for the electrochemical reaction (3) was given by Eq. (15). In general, the Nernst equation relates the activities (and/or fugacities) of the substances or species, a_i, in the cell's electrochemical reactions and the standard open-circuit potential, E^0, of the cell as:

$$E^{\mathrm{eq}} = E^0 - \left(\frac{RT}{nF}\right)\sum \nu_i \ln a_i \quad (17)$$

where ν_i is the stoichiometric number of the overall cell reaction which is positive for products and negative for reactants; n is the charge number of the cell reaction. Note that the current convention assumes that E^{eq} is a difference between the right-hand side and the left-hand side terminals. The standard value of the open-circuit potential is directly connected to the standard Gibbs energy, $\Delta_r G^0$, of the chemical reaction, which takes place in the electrochemical system:

$$\Delta_r G^0 = -nFE^0 \quad (18)$$

Because E^{eq} is an experimentally measurable value and E^0 can be obtained using an appropriate extrapolation, $\Delta_r G^0$ can be, in principle, experimentally derived using a high-temperature potentiometric system, a system where the electrochemical reactions are reversible and the electrode potential difference should be measured using a high-impedance electrometer. Moreover, if $\Delta_r G^0$ can be calculated using available thermodynamic data, the activities (and/or fugacities) of the chemical species can be experimentally derived. Therefore, the high-temperature potentiometric measurements can be a powerful tool in studying a number of thermodynamic phenomena if (1) the electrode electrochemical reactions are reversible and (2) the high-temperature potential measurements are reliable and reproducible within a few millivolts or less. It should be mentioned that, if the high-temperature thermodynamic properties (the standard Gibbs energies and the activity/fugacity coefficients) are available for all species of an electrochemical reaction, both E^0 and E^{eq} can be theoretically calculated and compared with observed potentiometric data. In this way, the reliability of an experimental potentiometric system can be confirmed.

For reliable high-temperature potentiometric studies, the measured open-circuit potential should be stable and reproducible within a few millivolts or less. Similar to the low-temperature electrochemical measurements, both the indicator and reference electrodes should be used in devising a high-temperature potentiometric cell. Both the electrochemical probes should be resistant to chemical degradation and pressure (mechanical) stress in high-temperature aqueous environment. The development of suitable indicator and reference electrodes that can reliably operate in aggressive high-temperature and high-pressure aqueous environment is still a challenge; this issue is addressed in Sects 10.2.1 and 10.2.2.

In high-temperature electrochemical kinetics studies, reactions at both electrodes are not at equilibrium, and there is consequently an overpotential $(E_i - E_i^{eq})$ on each of them. The potential difference measured between the two electrodes $(E - E^{eq})$ consists of the value of the two overpotentials, and it is not possible to determine the potential of the indicator electrode. Therefore, to obtain the potential value, a third so-called counter electrode must be used. In the three-electrode assembly, current flows between the indicator and the counter electrodes, and a high-impedance voltmeter is placed between the indicator and the reference electrodes to prevent any significant current flow through the reference electrode. In this case, the reference electrode shows a negligible overpotential and remains very close to its equilibrium value. In the described three-electrode device, the indicator electrode potential and its changes under electric current flow can be properly measured with respect to a reference potential, which should not be changed over the time of an electrochemical kinetics or corrosion study. In principle, the kind of material that is used for the counter electrode is not important, but one should make sure that a possible electrochemical degradation of the counter electrode will not contaminate the test solution. Also, the surface of the counter electrode should be sufficiently large to avoid a high current density at the electrode. Therefore, if one uses platinum foil or mesh for designing the counter electrode, no significant problem should occur. However, use of an appropriate reference electrode is probably the most significant challenge of the

electrochemical kinetics studies that are to be performed in high-temperature subcritical and supercritical aqueous solutions.

When current, I, passes through an electrochemical system electrolyte, an ohmic resistance, R, is observed between the electrodes inserted into the electrolyte. The IR drop is an ohmic voltage that originates from the electric current flowing in ionic solutions. In this case, the IR drop constitutes an unknown value and must be eliminated or minimized. A number of techniques can be applied to solve the undesirable problem of the IR drop.

One of the available approaches is to use the Luggin capillary, which is a tube made for high-temperature measurements from glass or Teflon at temperatures $<250\,^{\circ}C$ and ceramics at temperatures $>250\,^{\circ}C$. The end of the Luggin capillary is placed as close as possible to the test electrode surface, while the other end of the tube is electrically connected to another electrode compartment. If the Luggin capillary is filled with a high conductive electrolyte, the IR drop obviously decreases. Also, if a high-impedance voltmeter is used to prevent undesirable current flow through the Luggin capillary, the absence of current eliminates the IR drop and solves the problem in potentiometric studies.

Moreover, in the case of electrochemical kinetics and corrosion studies, some current should flow through the system and the IR drop is unavoidable and can be estimated using the current interruption technique [3].

Finally, in the case of the conductance studies, the IR drop is to be precisely measured and the measured conductivity of an electrochemical cell should be properly divided between the conductivity of solution and conductivity of other cell components including the electrode/solution interphases. Applying the electrochemical impedance spectroscopy can solve the problem of extracting the IR drop related only to aqueous solution from the measured total cell impedance [3].

10.2.1 High-temperature Reference Electrodes

As a minimum, a reference electrode must have a stable and reproducible potential during a course of measurement and the potential should not be changed if the test solution at the indicator electrode is varied. It is better if the reference electrode standard potential (with respect to the SHE scale) is thermodynamically defined. If the reference electrode solution is bridged into a test solution, then the diffusion (junction) potential between the test and reference solutions should be taken into account. A great deal of effort has been expended in recent years to develop a reliable and stable reference electrode that is suitable for measurements in high-temperature aqueous solutions. To date, two approaches have been employed for potentiometric measurements at elevated temperatures: (1) the use of an internal reference electrode operating within the high-temperature environment and (2) the use of an external reference electrode working at room temperature, but connected to the high-temperature environments by a nonisothermal electrolyte bridge. The first approach requires solving the well-known problem of the diffusion potential and the second approach additionally involves solving the problems of the thermal liquid junction and thermoelectric potentials. The nature of these and others undesirable potentials is considered in Sect. 10.2.3.

A number of electrochemical couples, such as Ag/AgCl, Ag/AgBr, Hg/Hg_2Cl_2, Hg/Hg_2SO_4, Ag/Ag_2SO_4, Hg/HgO, and

$Pb/PbSO_4$ have been tested for designing an internal high-temperature reference electrode. However, only the Ag/AgCl electrode [10] demonstrated the capability to be used at temperatures up to about 275 °C for a limited period of time. The reversible electrochemical reaction that takes place at the Ag/AgCl electrode is as follows: $AgCl(s) + e^- \leftrightarrow Ag(s) + Cl^-(aq)$ and the corresponding equation for the electrode potential, $E^{eq}_{Ag/Ag/Cl}$, can be constructed by using Eq. (16):

$$E^{eq}_{Ag/AgCl} = E^0_{Ag/AgCl} - \left(\frac{RT}{F}\right)\ln(a_{Cl^-}) \tag{19}$$

where a_{Cl^-} is the activity of the Cl^- (aq) ions and $E^0_{Ag/AgCl}$ is the standard electrode potential, which can be thermodynamically calculated if the standard Gibbs energies of formation, $\Delta_f G^0_i$, for all species involved in the electrochemical reaction are known:

$$E^0_{Ag/AgCl} = \frac{-\Delta_r G^0_{Ag/AgCl}}{F} = -\frac{\left(\Delta_f G^0_{Ag} + \Delta_f G^0_{Cl^-} - \Delta_f G^0_{AgCl}\right)}{F} \tag{20}$$

While all necessary thermodynamic properties for calculating $E^0_{Ag/AgCl}$ are available, no reliable high-temperature internal reference electrode has been developed for a temperature range above 300 °C, apparently because of a chemical degradation process of the Ag/AgCl electrochemical couple in a hydrothermal environment. Therefore, no reliable studies have been carried out to find a suitable internal reference electrode that can be employed in high-temperature subcritical and supercritical aqueous solutions.

The chemical stability problem of the internal reference electrode was resolved by the development of the external (pressure-balanced) reference electrode, in which the electrochemical couple (e.g. Ag/AgCl) is maintained at ambient temperature and is connected to the high-temperature zone by a nonisothermal electrolyte bridge [11]. While this approach overcomes the stability and chemical degradation problems, it introduces the problem of the Soret effect. The Soret effect arises from a coupling between irreversible processes of the heat and mass transport along the nonisothermal electrolyte bridge and generates an additional thermal diffusion potential (a function of temperature, composition, and time) that can be substantial in magnitude. One of the latest designs of the external reference electrode applies a flow-through technique [12], which minimizes interference from contaminates (e.g. corrosion products) and from the Soret effect, thereby providing a stable and highly accurate reference potential within a couple of millivolts or less. However, an additional streaming potential is generated if an electrolyte solution flows through a tiny capillary channel. This potential should also be taken into account.

The flow-through external reference electrode can be used at any desirable temperature and pressure (until the aqueous solution is sufficiently conductive), and the stability of the electrode potential does not depend on the state parameters. This statement has experimentally been confirmed at temperatures from 25 to 400 °C and pressures up to 350 bar. The specific feature of the flow-through external reference electrode is that the reference solution flows through the electrode at a constant velocity. Therefore, the concentration of solution across the thermal

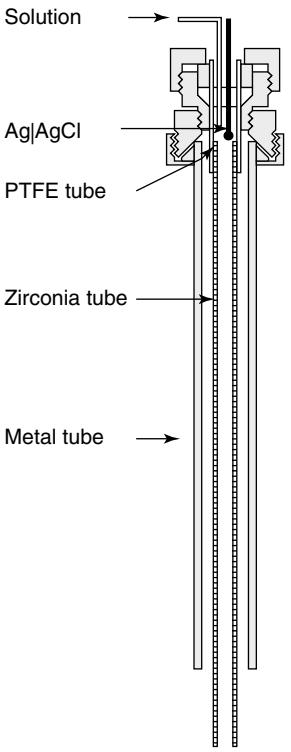


Fig. 1 Flow-through external reference electrode[12, 13].

junction remains constant and any uncertainty in the thermal diffusion potential can be eliminated for measurements at a given temperature and pressure. If desirable, the thermal liquid junction potential can be either calculated or experimentally estimated and, in this way, the electrode potential value can be evaluated with respect to the SHE scale. In the studies carried out so far, the flow-through reference electrode was used in a thermocell, which has either platinum or yttria-stabilized zirconia (YSZ) indicator electrodes as pH sensors for potentiometric measurements up to 400 °C [13]. In these tests, 0.1 mol kg^{-1} NaCl aqueous solution was used as a nonisothermal electrolyte bridge (reference electrode solution) to electrically connect the Ag/AgCl electrochemical couple, kept at ambient temperature and high-temperature aqueous environment, where an indicator electrode is located. The potential of the flow-through external reference electrode was found to be stable within 1–3 mV as long as the reference electrode solution pumps through, using a high-pressure chromatography pump. The need to use a high-pressure pump for operating the flow-through reference electrode is a disadvantage of this technique. A schematic of the extensively tested flow-through external reference electrode is presented in Fig. 1.

In conclusion, while a reliable flow-through external reference electrode has been successfully developed, there is still the pressing need for the development of a robust internal reference electrode, which could reliably operate in high-temperature subcritical and supercritical aqueous solutions.

10.2.2 High-temperature Indicator Electrodes

The indicator electrode must have a stable and reproducible potential for a course of measurement and should be able to respond in a Nernstian manner to varying conditions in the high-temperature aqueous environment. In other words, the activity of the dissolved species, a_i, and the standard open-circuit potential, E^0, should be, in principle, definable by measuring the open-circuit potential between the indicator and reference electrodes and applying the Nernst equation (17).

While a number of the indicator electrodes have been tested for operation over a wide range of temperatures, only the

platinum/hydrogen, Pt(H_2), and the YSZ electrode with mercury/mercury oxide electrochemical couple, YSZ(Hg/HgO), were found to be capable of operating in the Nernstian manner at temperatures up to 400 °C. The corresponding electrochemical reversible reactions that take place at the Pt(H_2) and YSZ(Hg/HgO) electrodes are, respectively, $2H^+(aq) + 2e^- \leftrightarrow H_2(g)$ and $2H^+(aq) + 2e^- + HgO(s) \leftrightarrow Hg(l) + H_2O(l)$. The corresponding equations for the electrode potentials of the Pt(H_2) and YSZ(Hg/HgO) electrodes can be written using Eq. (16) as follows:

$$E^{eq}_{Pt(H_2)} = \left(\frac{RT}{2F}\right) \ln\left(\frac{a^2_{H^+}}{f_{H_2}}\right) \quad (21)$$

$$E^{eq}_{YSZ(Hg/HgO)} = E^0_{YSZ(Hg/HgO)} + \left(\frac{RT}{2F}\right) \ln\left(\frac{a^2_{H^+}}{a_{H_2O}}\right) \quad (22)$$

where $E^0_{YSZ(Hg/HgO)}$ is the standard YSZ(Hg/HgO) electrode potential, which does not depend on the solution composition but depends on the standard Gibbs energies of formation, $\Delta_f G^0_i$, of the species involved in the electrochemical reaction [H^+ (aq), HgO(s), Hg(l), and H_2O(l)]:

$$E^0_{YSZ(Hg/HgO)} = \frac{-\Delta_r G^0_{YSZ(Hg/HgO)}}{2F} = \frac{-\begin{pmatrix} \Delta_f G^0_{H_2O} + \Delta_f G^0_{Hg} \\ -\Delta_f G^0_{HgO} - 2\Delta_f G^0_{H^+} \end{pmatrix}}{2F} \quad (23)$$

Note that $\Delta_f G^0_{H^+}$ and $\Delta_f G^0_{H_2}$ are zero at any temperature, so $E^0_{Pt(H_2)} = 0$ and therefore does not come up in Eq. (21). Also, it should be mentioned that $E^0_{YSZ(Hg/HgO)}$ does not depend on the YSZ membrane properties and the only requirement is a sufficient O^{2-} conductivity at the temperature of interest. It was experimentally found that when using a modern high-impedance ($>10^{14}\Omega$) electrometer the lowest temperature for using an YSZ membrane can be around 100 °C. Therefore, the YSZ(Hg/HgO) electrode cannot be practically used at temperatures below 100 °C.

The Pt(H_2) indicator electrode has been widely used in measuring pH in the concentration cells at temperatures below 300 °C [2, 14] and in the flow-through electrochemical systems at temperatures up to 400 °C [13]. The schematic of the flow-through Pt(H_2) electrode, which has been used together with the flow-through reference electrode described above, is given in Fig. 2 [13]. The YSZ(Hg/HgO) indicator

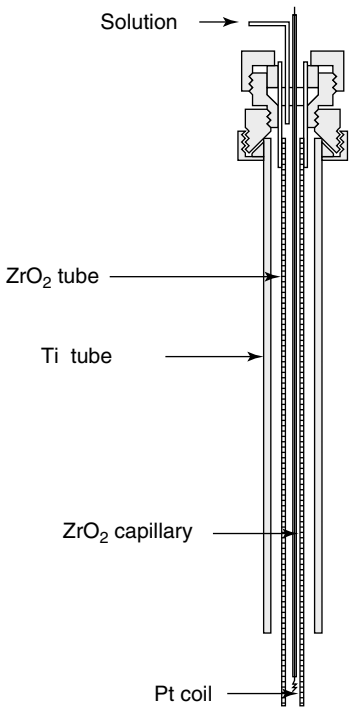


Fig. 2 Flow-through Pt(H_2) electrode [13].

electrode has also been used in both the static systems [15] at temperatures up to 450 °C and flow-through electrochemical cells at temperatures up to about 400 °C [16, 17]. Both electrodes were found to be Nernstian and capable of measuring the electrode potential to a precision of less than ±5 mV. The main disadvantage of the Pt(H_2) electrode is that the reversibility of the electrode can be significantly damaged by certain "poisons" that interfere with the normal operation of the hydrogen electrode. Also, the fugacity of H_2 should be well established for estimating the activity of H^+ (aq) using Eq. (21). The main disadvantage of the YSZ(Hg/HgO) electrode is the relative complexity of the design. Also, the activity of H_2O should be known for calculating the activity of H^+ (aq) using Eq. (22). In addition, the chemical stability of the commercially available membrane is not still sufficient if the electrode is to be used in an aggressive, high-temperature, aqueous environment [16]. Note that other electrochemical couples (e.g. Cu/CuO) have been tested to be incorporated into the YSZ indicator electrode but have not demonstrated the necessary chemical stability as the Hg/HgO electrochemical couple did. The design of the flow-through YSZ(Hg/HgO) indicator electrode, which has been used together

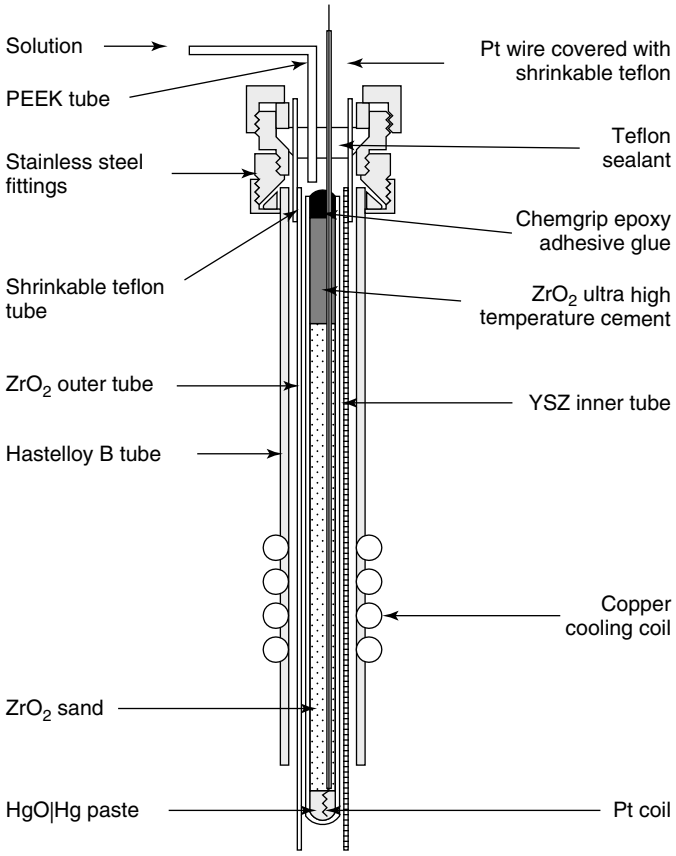


Fig. 3 Flow-through YSZ(Hg/HgO) electrode [16, 18].

with the above-described flow-through reference electrode, is given in Fig. 3 [16, 18].

The usefulness of a number of metal/metal-oxide (e.g. Ir/IrO_2, Zr/ZrO_2, W/WO_2, etc.) electrodes and the glass electrode has been tested over a wide range of temperatures. However, the existence of the Nernstian behavior has not been well demonstrated yet. The glass electrode can probably be employed at temperatures up to about 200 °C but was found to be impractical owing to an inconvenient design for high-temperature subcritical and supercritical aqueous solutions.

10.2.3 Undesirable Potentials

The diffusion (E_D), thermal diffusion (E_{TD}), thermoelectric (E_{TE}), and streaming (E_{STR}) potentials are the potentials that are not desirable in the potentiometric measurements but can occur. If these phenomena have an effect, they should be taken into account to correctly translate the measured open-circuit potential to thermodynamic properties of the electrochemical system using a generalized Nernst equation:

$$E = E^0 - \left(\frac{RT}{nF}\right)\sum \nu_i \ln a_i + (E_D + E_{TD} + E_{TE} + E_{STR}) \quad (24)$$

Note that the above-mentioned phenomena are irreversible in nature and can be properly understood on the basis of the linear irreversible thermodynamics. The diffusion potential, E_D, arises if two solutions are in contact. This phenomenon is a result of the different mobility properties of the ion species and can theoretically be estimated if the individual ion electric conductivities and activities of the species are known. The diffusion potential should be either minimized or calculated if the necessary properties of the species are available. Apparently, E_D can be a small value in high subcritical and supercritical aqueous solutions. However, this assumption should be carefully confirmed before applying it to a particular high-temperature liquid junction. At temperatures below 350 °C, if E_D is not minimized, it could be as substantial as 30 mV or larger [12]. Reliable estimation of the diffusion potential is a serious challenge in high-temperature electrochemical studies.

The thermal diffusion potential, E_{TD}, arises if an electrochemical system is nonisothermal. This phenomenon is due to the heat transport of ionic species and can be taken into account if the individual ion entropy of transport, conductivity, and activity coefficients of the species of interest are known. Therefore, the thermal diffusion potential depends on the temperature, pressure, and composition of the electrolyte liquid junction. Also, E_{TD} is a function of the temperature gradient and can be a substantial value from tens to hundreds of millivolts [19].

The thermoelectric potential, E_{TE}, is due to the heat transport of the electrons and arises if an electron conductor (usually a wire) is in nonisothermal condition. E_{TE} is a function of the temperature gradient and, for the most common wire materials, is usually up to a few millivolts. E_{TE} can be calculated over a wide range of temperatures for most of the wire materials such as Pt, Ag, Cu, Fe, Ni, and so on [13].

The electric potential which is set up in a capillary channel (or porous material) at zero electric current, owing to a material flow of solution, is called the *streaming potential*, E_{STR}. The streaming potential depends on (1) capillary channel material, (2) composition of flowing

solution, (3) solution flow rate, and (4) temperature. Presently, there are no data to theoretically evaluate E_{STR} at high temperatures. However, if the flow rate is sufficiently low there is a linear dependency between the streaming potential and flow rate and this linear dependence can be used to eliminate E_{STR} by extrapolating the measured potentials to zero flow rate [13].

10.2.4 High-temperature Reference and Buffer Solutions

For assessing the viability and accuracy of high-temperature potentiometric measurements, the reference systems should be used. If Pt(H_2) or YSZ(Hg/HgO) electrodes are used as the indicator electrodes, an aqueous solution with well-known activity of H^+ (aq) should be used as the reference systems. At temperatures below 250 °C, the dilute aqueous solutions of strong acids and bases, such as HCl(aq) or NaOH(aq), can be employed to precisely calculate the activity of H^+ (aq) so that the measured potential can be compared with the calculated one within a few millivolts or less. If HCl(aq) or NaOH(aq) is to be used at temperatures above 250 °C, the association constants of the electrolytes should be taken into account. Furthermore, at these temperatures, the precision of the calculated activities of H^+ (aq) can be decreased. However, even in the low-density, supercritical aqueous solution, a reference system, which consists of a couple of three-component aqueous solutions, can be found to test the accuracy of the Pt(H_2) or YSZ(Hg/HgO) electrodes within about ±3 mV. Each of the three-component aqueous solutions consists of NaCl and either HCl or NaOH such that the concentration of NaCl is significantly greater than the concentration of another solute (such as HCl or NaOH). The approach was theoretically and experimentally confirmed at temperatures up to 400 °C and densities down to 0.17 g cm^{-3} and does not require any knowledge of the association constants [20].

The buffer solutions play a significant role in low-temperature potentiometric and electrochemical kinetics measurements. It is highly desirable to estimate a set of the buffer solutions that can be used at temperatures above 100 °C. Thus far, little has been done for developing a necessary set of high-temperature buffer systems. However, aqueous 0.05 mol kg^{-1} potassium hydrogen phthalate solution has been shown as an appropriate buffer system to be used at temperatures up to about 225 °C [21].

10.3 High-temperature Potentiometry

During the last 30 years, Oak Ridge National Laboratory (ORNL) and others have used the hydrogen-electrode concentration cells (HECC) [14] for potentiometric measurements of homogeneous [22] and heterogeneous [23] protolytic aqueous systems to 300 °C over a wide range of ionic strength. Many acid–base, metal ion hydrolysis, metal complexation, metal-oxide solubility, and surface adsorption reactions have been studied using this technique [22, 23]. The most recent design of the stirred HECC is described in Ref. 14.

While the electrochemical experimental techniques used at temperatures below 300 °C are out of the scope of this chapter, a brief description of the HECC will be given below, owing to the significance

of this system [2, 14]. The HECC pressure vessel is usually machined from Hastelloy B, which was chosen for its resistance to corrosion, especially stress cracking in the presence of chloride, and hydrogen embrittlement. Two concentric internal cups are machined from Teflon with the smaller, inner compartment containing a reference solution whose acidity or basicity is fixed by a known ''stoichiometric'' amount of a strong acid or strong base, or buffer. The H_2(g) pressure is the same throughout the cell so that the pH measurements do not require knowledge of the partial pressure of hydrogen, which is fixed by conducting at least five pressurizations and release cycles at room temperature to purge the cell of free oxygen. The final pressure of H_2(g) is typically 1–3 MPa. The cell is then heated in an oil bath or furnace to the desired temperature while the solutions in both compartments are stirred magnetically. A liquid junction is maintained between the two solutions using a small porous plug of Teflon compressed into a hole in the bottom of the reference cup. Titrations are conducted by injection of a titrant through PEEK (polyetheretherketone) (at low temperature) and platinum lines (at high temperature) from a calibrated Zircalloy positive-displacement pump. The electrochemical diagram of the HECC is as follows:

$$\begin{aligned} &\mathrm{Cu|Pt(H_2)|Reference\ H^+(aq)/H_2(aq)} \\ &\quad \mathrm{Solution\ 1||Indicator\ H^+(aq)/H_2(aq)} \\ &\quad \mathrm{Solution\ 2|Pt(H_2)|Cu} \end{aligned} \tag{25}$$

where a single vertical bar represents a phase boundary and a double vertical bar represents the liquid junction between miscible liquids. Using the generalized Nernst equation (24), the open-circuit potential, E_{HECC}, can be presented as follows:

$$E_{\mathrm{HECC}} = \left(\frac{RT}{F}\right) \ln\left[\frac{(a_{\mathrm{H^+}})_2}{(a_{\mathrm{H^+}})_1}\right] + E_{\mathrm{D}} \tag{26}$$

where $a_{\mathrm{H^+}}$ is the activity of the hydrogen ion related to the molal concentration scale. In Eq. (26), the subscripts 1 and 2 symbolize the reference and indicator electrode solutions, respectively, and E_{D} is the liquid junction potential to be theoretically calculated. Fundamental to the operation of the HECC is that both solutions be maintained at the same ionic strength by the addition of a strong ''supporting or inert'' electrolyte, such as NaCl or NaF_3CSO_3. This is crucial to the assumption that the activity coefficient of the hydrogen ions is the same in both compartments of the cell and the value of E_{D} is minimized by the predominance of the same ions on both sides of the junction. The precision of potentiometric measurements carried out using the HECC system is generally less than ± 0.1 V for well-behaved electrolyte solutions.

The hydrogen concentration cells described above cannot be used at temperatures above 300 °C because Teflon loses its thermal stability. However, a flow-through potentiometric system can be used [13, 16] and allows the high-temperature potentiometric measurements in high-temperature subcritical and supercritical aqueous solutions. A schematic of the electrochemical cell is shown in Fig. 4 [16]. The cell is machined using a corrosion-resistant alloy and has four ports into which different components can be sealed for use at high pressures and temperatures. The components include a flow-through external Ag/AgCl reference electrode, a flow-through Pt(H_2) indicator electrode, a flow-through YSZ(Hg/HgO)

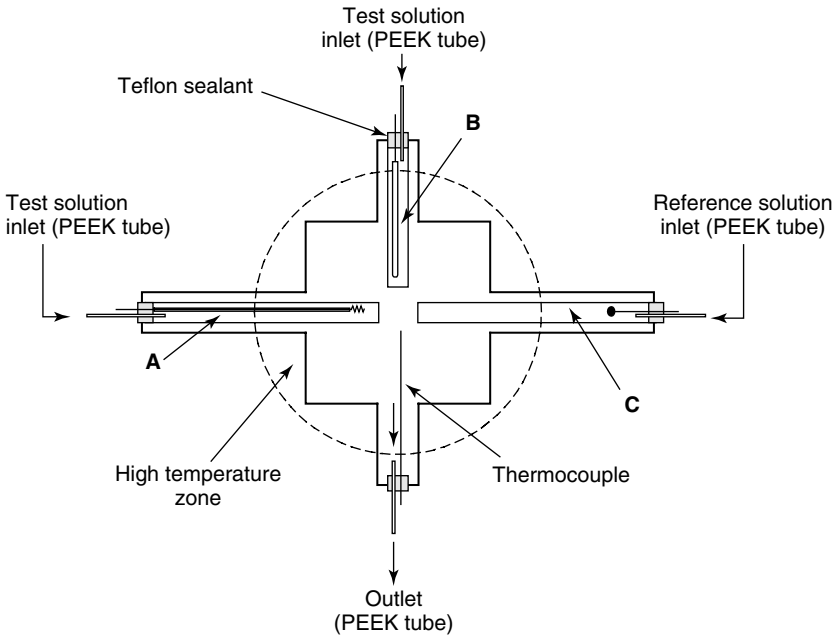


Fig. 4 The flow-through electrochemical cell for potentiometric measurements at temperatures up to 400 °C: A – flow-through Pt(H_2) electrode, B – flow-through YSZ(Hg/HgO), C – flow-through external Ag/AgCl reference electrode [13, 16].

indicator electrode, and a thermocouple. The design has a four-way, once-through, pumped fluid circulation system that can pump fluid through the electrodes at rates faster than thermal diffusion so that no concentration gradients result from the Soret effect. Only the sensing portion of the system needs to be, and is, maintained at controlled temperatures and pressures. The purity and concentration of the solutions are maintained by relatively rapid flow. Therefore, contamination is minimized by this flow, and hence, dissolution and corrosion of the system are significantly reduced. Because the low-temperature input flow is only in contact with glass, Teflon, and PEEK tubes (which all have high corrosion resistance) and the high-temperature outflow is only in contact with zirconia and platinum, significant control is achieved on the solution composition at the sensing portion of the system. Use of the above-mentioned materials is a crucial requirement for accurate high-temperature potentiometric measurements. The achieved precision of the potentiometric measurements, using the flow-through electrochemical cell, is found to be ± 5 mV or less. As an example, employing the system described above, the association constants of HCl(aq) at temperatures from 300 to 400 °C were recently obtained [16, 17]. Also, using both the Pt(H_2) and YSZ(Hg/HgO) electrodes the Henry's constant of H_2(aq) was obtained at temperatures between 300 and 450 °C [15, 24]. Therefore, it was found that reliable potentiometric data could be obtained using an YSZ (Hg/HgO) electrode at temperatures up to 450 °C and densities down to 0.1 g cm^{-3}. An electrochemical diagram of the nonisothermal electrochemical system

(thermocell) described above, which consists of the flow-through external Ag/AgCl reference electrode and a flow-through YSZ(Hg/HgO) electrode, can be presented as follows:

$$(T_1)\text{Cu}|\text{Ag}|\text{AgCl}|\text{Reference Electrode}|\text{Reference Solution}\ldots(T_2)\text{Reference Solution}\|\text{Indicator Solution}|\text{Indicator Electrode}|\text{YSZ}|\text{HgO}|\text{Hg}|\text{Pt}\ldots(T_1)\text{Pt}|\text{Cu} \quad (27)$$

where T_1 is the ambient temperature and T_2 is an elevated temperature higher than T_1. Note that Cu in Eq. (27) represents the wires connecting the terminals of the system to an electrometer. The open-circuit potential of the thermocell, E_{TC}, can be expressed as:

$$E_{\text{TC}} = [E^0_{\text{YSZ(Hg/HgO)}}]_{T_2} - [E^0_{\text{Ag/AgCl}}]_{T_1} + \left(\frac{RT_2}{F}\right)\ln[(a_{\text{H}^+})_{T_2}(a^{-0.5}_{\text{H}_2\text{O}})_{T_2}] + \left(\frac{RT_1}{F}\right)\ln[(a_{\text{Cl}^-})_{T_1}] + E_{\text{D}} + E_{\text{TD}} + E_{\text{TE}} + E_{\text{STR}} \quad (28)$$

where the subscripts T_1 and T_2 symbolize the ambient and elevated temperatures respectively. Note that one of the disadvantages of the flow-through electrochemical system is the existence of the streaming potential, which should be taken into account for accurate potentiometric studies. To eliminate E_{STR}, it is necessary to extrapolate the measured open-circuit potentials to zero flow rate. Therefore, at least four experimental points should be measured for providing a reliable linear extrapolation.

A flow-through thermocell, similar to that described in Refs 13, 16, was recently reported in Ref. 25. The reference Pt(H_2) electrode was maintained at ambient temperature with the flow of both reference and test solutions coming from opposite ends of the cell past each electrode. The authors used the flow-through thermocell to measure the acid dissociation constant for phenol to 400 °C and 34.9 MPa [26]. Hence, it was firmly proved that the flow-through potentiometric technique could be used at temperatures up to 400 °C and densities down to 0.2 g m^{-3}.

10.4 High-temperature pH Measurements

Perhaps the most important parameter in defining the chemical properties of high-temperature solutions is the $\text{pH} = -\lg(a_{\text{H}^+})$, which is of critical importance in understanding such high-temperature processes as acid–base equilibria, metal ion hydrolysis, metal complexation, salt solubility, and so on.

The definition of pH by the International Union of Pure and Applied Chemistry (IUPAC) is as follows:

$$\text{pH} = -\lg(m_{\text{H}^+}\gamma_{\text{H}^+}) \quad (29)$$

where m_{H^+} and γ_{H^+} are, respectively, the molal concentration and activity coefficient of the H^+(aq) ion. Because γ_{H^+} is an unmeasurable quantity, pH is defined operationally in terms of the method used to measure it by means of the electrochemical cell. The IUPAC recommendation [8] is to use the following cell for pH measurements:

$$\text{Cu}|\text{Reference Electrode}|\text{KCl Solution}\ (>3.5\ \text{mol kg}^{-1})\|\text{H}^+(\text{aq})\text{Solution } S|\text{H}_2|\text{Pt}|\text{Cu} \quad (30)$$

which can be used at temperatures below 100 °C. Solution S in Eq. (30) is a standard reference solution or a buffer solution. At temperatures from 0 to 95 °C, 22 standard reference solutions were recommended by IUPAC [8]. Each of the solutions is assigned to a specific pH value at a given temperature between 0 and 95 °C. The pH of an arbitrary solution X, denoted by pH(X), is defined in relation to the pH of a standard reference solution S, denoted by pH(S) as:

$$\mathrm{pH}(X) - \mathrm{pH}(S) = -\frac{E(X) - E(S)}{(RT/F)\ln 10} \quad (31)$$

where $E(S)$ is the open-circuit potential of the cell (30) and $E(X)$ is the open-circuit potential of a similar cell as follows:

$$\begin{aligned} &\mathrm{Cu|Reference\ Electrode|KCl\ Solution} \\ &(>3.5\ \mathrm{mol\ kg^{-1}})||\mathrm{H^+(aq)Solution} \\ &X|\mathrm{H_2|Pt|Cu} \end{aligned} \quad (32)$$

In practice, a glass electrode is almost always used in place of the Pt(H_2) electrode. A glass electrode has a deviation from the H^+ (aq) ion response function (non-Nernstian behavior) and, therefore, should be calibrated using a set of the standard (buffer) solutions.

While the above approach works very well at temperatures below 100 °C, it is difficult to apply the IUPAC recommendations at temperatures above 100 °C when a high-temperature system should be pressurized. Definitely, at temperatures below 300 °C the HECC represented by (25) can be employed for pH measurements in the solutions where the half-reaction of the Pt(H_2) electrode is a reversible process. Accuracy of the measurements could be ±0.01 pH units and is mainly limited from estimating the diffusion potential in Eq. (26).

At temperatures from 100 to 400 °C, the flow-through electrochemical system described in Refs 13, 16 was demonstrated as an appropriate device for high-temperature pH measurements achieving an accuracy of about ±0.05 pH units. However, this level of precision can be obtained only if an extrapolation to zero flow rate is carried out. Therefore, a few experimental points should be picked up for a reliable extrapolation, and this approach is a time-consuming procedure. Another approach, which was recently developed using the same device for a faster high-temperature pH measurement, introduces a calibration coefficient, α, which takes into account the irreversible thermodynamic contributions that are presented in Eq. (24). If a flow-through YSZ(Hg/HgO) probe is used as a pH-sensing indicator electrode, the pH value of a test solution X can be evaluated using the following expression [27]:

$$\mathrm{pH}(X) - \mathrm{pH}(S) = -\alpha\frac{E(X) - E(S)}{2.303RT/F} - \frac{1}{2}\lg\left(\frac{a_{\mathrm{H_2O}}(S)}{a_{\mathrm{H_2O}}(X)}\right) + \frac{E_\mathrm{d}(X) - E_\mathrm{d}(S)}{2.303RT/F} \quad (33)$$

where S denotes the standard solution, $E(X)$ and $E(S)$ are, respectively, the open-circuit potential for solutions X and S, $a_{\mathrm{H_2O}}(X)$ and $a_{\mathrm{H_2O}}(S)$ are, respectively, activities of water of solutions X and S, and $E_\mathrm{d}(X)$ and $E_\mathrm{d}(S)$ are, respectively, the diffusion potentials of solutions X and S with respect to the reference electrode potential. The calibration coefficient, α, should be estimated in a range of pH of interest using reference systems or buffer solutions as described in Sect. 10.2.4. The approach has been tested so far for a temperature range up to about 300 °C [27] and additional efforts are needed to find

out if this method can be used at higher temperatures. The expected accuracy of the pH measurements is around ±0.1 pH units.

10.5 High-temperature Electrokinetic Studies

Four well-known electrokinetic phenomena that can be experimentally studied are (1) electroosmosis – if there is movement of a liquid along a solid phase in an electric field, (2) electrophoresis – if there is movement of a solid suspended in a liquid in an electric field, (3) streaming potential – if a potential difference is formed between the upstream and the downstream ends of a liquid flow in either capillary tubes or porous plugs, and (4) sedimentation potential – if a potential is formed owing to the falling of a suspended solid phase in a liquid phase. Only the electrophoresis and streaming potential measurements have been applied so far for high-temperature and high-pressure electrochemical measurements. Using the electrokinetic approaches mentioned above, zeta potentials can be derived. This is usually defined as the potential of the electric double layer at the slipping plane between the bulk solvent and relatively stagnant layer close to the solid surface. Zeta potential generally depends on temperature and pH and is usually positive in highly acidic solutions and negative in highly basic solutions. The pH value at which zeta potential equals zero is defined as the isoelectric point (IEP) and an important characteristic of a metal-oxide/water interface. The flow-through system described in Ref. 13 was employed to attempt obtaining zeta potentials at the zirconia/aqueous solution interface at temperatures up to 400 °C, but a relatively large experimental error did not allow the reliable estimation of the zeta potential temperature dependence. Most recently, a successful effort was made to develop a high-temperature microelectrophoresis system [28]. A limited amount of high-temperature zeta potential data have been thus far obtained for ZrO_2/water and TiO_2/water interfaces at temperatures up to 200 °C [28, 29]. A significant effort is needed to extend the temperature range of operating the above-described microelectrophoresis system to a supercritical region and to carry out the zeta potential measurements at temperatures up to 400 °C.

10.6 High-temperature Electrochemical Kinetics and Corrosion

A systematic understanding of the electrochemical kinetics and mechanisms of electrochemical reactions, including corrosion processes, requires the measurements of the appropriate electrochemical kinetic parameters in Eqs (13) and (14). In particular, there is a significant scarcity of the key electrochemical kinetic parameters, the exchange current densities (j_o) and the reduction (α_{red}) and oxidation (α_{oxid}) transfer coefficients, for charge-transfer reactions. It turns out that even the most important HER and oxygen electrode reaction (OER) have rarely been systematically studied at temperatures above 100 °C. The reason for this state of research is not difficult to determine – the electrochemical kinetic measurements are currently not easy to perform at elevated temperatures and few of the available electrochemical sensors and/or systems are now available for performing high-quality electrochemical kinetics and corrosion experiments. Most of the electrochemical

kinetics reactions and corrosion processes are heterogeneous processes, and obtained experimental data reflect both the mass-transfer and charge-transfer phenomena. However, the above-mentioned parameters (j_o, α_{red}, α_{oxid}) are strictly charge transfer in nature, and it is necessary to delineate these phenomena in any electrochemical kinetics or corrosion experiment. To solve the problem, one needs to carry out the electrochemical kinetics or corrosion studies under well-defined hydrodynamic conditions. Nevertheless, the traditional rotating-disk (ring) electrochemical system, which has been widely used at low temperatures, has not been well implemented in high-temperature electrochemical studies yet. Thus far, only a few controlled hydrodynamic systems for the measurement of electrochemical kinetic parameters have been tested at temperatures below 300 °C (where Teflon is still mechanically stable). Although a recent study of the HER on mercury in HCl(aq) solutions has been carried out at temperatures up to 300 °C [30], it will be difficult to extend the temperature range beyond the temperature limit. In addition, a high-temperature wall-tube electrode cell for electrochemical kinetic studies at temperatures up to 200 °C has been described in Ref. 31. This technique could probably be used for electrochemical kinetics studies at higher temperatures, but has not been tested yet at temperatures above 200 °C. Another high-temperature tubular, flow-through electrochemical system has been designed to operate at temperatures up to about 250 °C [32, 33], and this approach looks promising for use in high-temperature subcritical and supercritical aqueous systems.

It turns out that so far only very few electrochemical kinetics studies have been carried out at temperatures above 300 °C. In Ref. 34, the authors reported a study in which well-defined steady-state cyclic voltammetry was applied to learn a number of electrochemical reactions in near-critical and supercritical fluids at temperatures up to 385 °C [34].

In another study [35], the electrochemical emission spectroscopy (electrochemical noise) was implemented at temperatures up to 390 °C. It is well known that the electrochemical systems demonstrate apparently random fluctuations in current and potential around their open-circuit values, and these current and potential noise signals contain valuable electrochemical kinetics information. The value of this technique lies in its simplicity and, therefore, it can be considered for high-temperature implementation. The approach requires no reference electrode but instead employs two identical electrodes of the metal or alloy under study. Also, in the same study electrochemical noise sensors have been shown in Ref. 35 to measure electrochemical kinetics and corrosion rates in subcritical and supercritical hydrothermal systems. Moreover, the instrument shown in Fig. 5 has been tested in flowing aqueous solutions at temperatures ranging from 150 to 390 °C and pressure of 25 MPa. It turns out that the rate of the electrochemical reaction, in principle, can be estimated in hydrothermal systems by simultaneously measuring the coupled electrochemical noise potential and current. Although the electrochemical noise analysis has yet to be rendered quantitative, in the sense that a determination relationship between the experimentally measured noise and the rate of the electrochemical reaction has not been finally established, the results obtained thus far [35] demonstrate that this method is an effective tool for

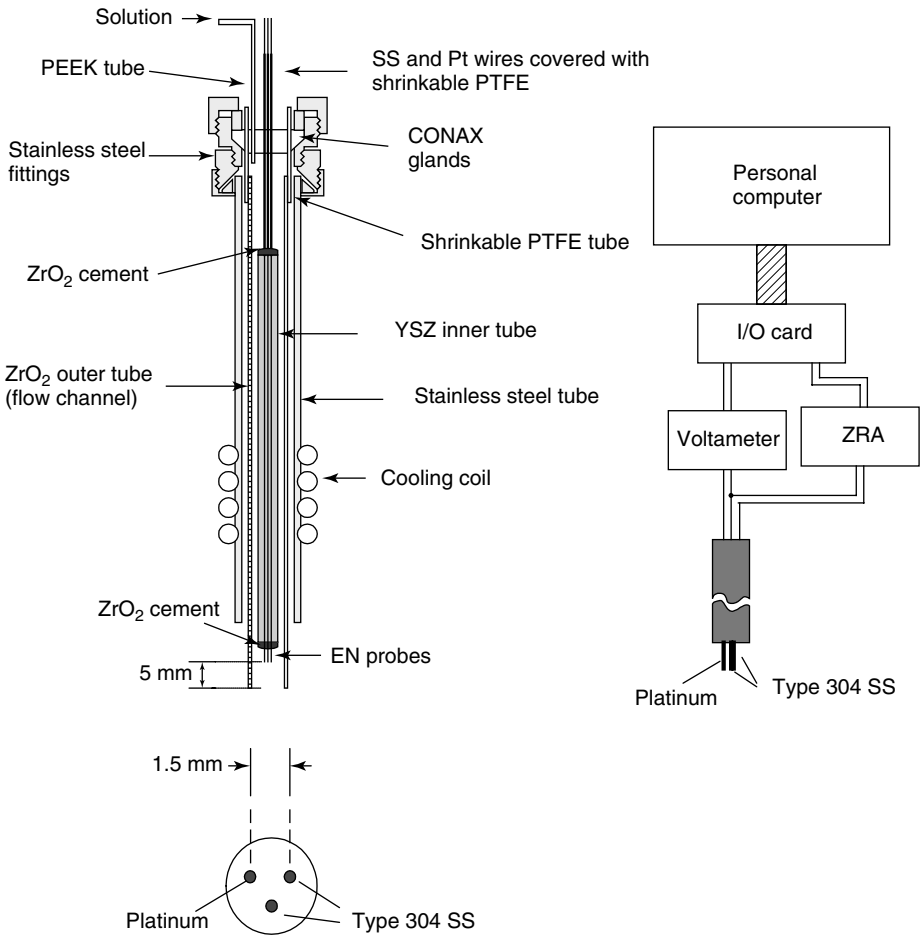


Fig. 5 Schematic of the electrochemical emission spectroscopy system [35].

studying the electrochemical kinetics phenomena in high-temperature subcritical and supercritical aqueous solutions.

In conclusion, it should be mentioned that the electrochemical kinetics of high-temperature subcritical and supercritical aqueous systems is an underdeveloped field of research, and a significant effort is still needed to make this a reliable and well-understood approach. Also, in regard to the electrochemical emission spectroscopy method to be used in supercritical conditions, a basic understanding of the phenomenon of electrochemical noise is still required before such information can be applied to practical systems. To our knowledge, no reliable electrochemical corrosion technique has been developed thus far, and it is obvious that little progress can be achieved in studying the electrochemically driven corrosion processes that occur in high-temperature subcritical and supercritical aqueous systems until a well-defined electrochemical kinetics instrument is finally developed.

10.7 High-temperature Electrical Conductance Techniques

The ionic conductivity measurements in aqueous solutions under ambient conditions are relatively simple. However, in order to extend the conductivity tests to a high-temperature region, specialized techniques are required. A detailed description of the available techniques for the conductance studies in high-temperature subcritical and supercritical aqueous solutions was recently reviewed in detail [22, 36]. The experimental challenges of high-temperature conductivity measurements include the need to use noncorrosive materials, accurately control temperature and pressure, and use a thermally stable and chemically inert isolating material. The first experimental conductivity systems had a static design and did not allow the study of high-temperature aqueous solutions both at low concentrations ($<10^{-5}$ mol kg^{-1}) and close to the critical point of water. However, accurate conductance values for dilute solutions are essential if one wants to calculate the ion association constants. Also, the physical chemistry of aqueous solutions near critical point is an important area of research for a number of applications. Thus, to overcome the static design challenges, a flow-through conductance system for high-temperature applications has been recently developed. The modest system design is described in Ref. 37. The flow-through cell allows rapid and accurate electric conductance measurements in high-temperature subcritical and supercritical aqueous solutions with concentration as low as 10^{-8} mol kg^{-1}, and even in the vicinity of the critical point of water. The cell is constructed from a platinum/rhodium cup, an outer electrode, and gold soldered to platinum/rhodium tubing, which is used as an inlet tube (Fig. 6). At the edge of the cup an annealed gold washer is located. The washer is on the top of a sapphire disc insulator, through which is connected the inner electrode and a platinum/rhodium tube. The inner electrode tube is gold-filled at one end. Two small holes on the other end act as the solution outlet. A particular feature of the cell is that the permanent solution flow sweeps the contaminants dissolving from the sapphire insulator out of the measuring zone and, thus, minimizing the adsorption of undesirable species on the cell surface. In tests, a significant improvement in speed and accuracy was achieved by employing

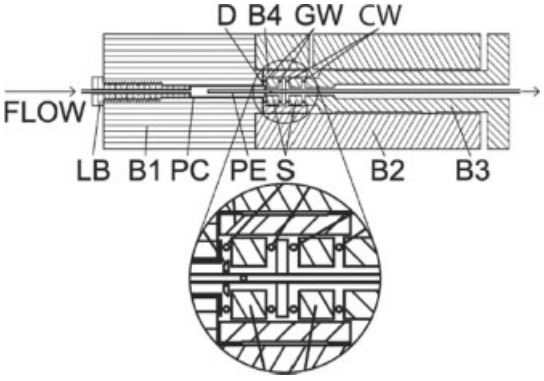


Fig. 6 Cross section of a high-precision conductivity cell [37]: B1, B2, and B3 are cylindrical titanium blocks; B4 is a titanium sleeve; PC is the platinum cup; D is the diamond shield; S are the sapphire insulators; H is a hole for the solution exit; GW are the gold washers; CW are the carbon (graphite) washers; PE is the platinum electrode; and LB is a titanium screw.

this cell design, and the reported precision was between 0.1 and 1%, depending on the electrolyte concentration and solution density.

10.8 Conclusions

Advances in electrochemical techniques for studying high-temperature subcritical and supercritical aqueous solutions with emphasis on new cell configurations and electrodes have been reviewed in this chapter.

Whereas the standard electrode potentials of many half-cell reactions have been known at ambient conditions and can be easily found in a number of reference books, almost none of them are documented for a region of high-temperature subcritical and supercritical conditions. Therefore, the creation of well-established approaches for developing a comprehensive list of the standard potentials measured over a wide range of temperatures remains a challenge for high-temperature experimental electrochemistry. The recently developed instruments for potentiometric studies at temperatures above 300 °C can be useful for developing such a database.

Although a suitable external Ag/AgCl reference electrode has been successfully developed for operating in high-temperature subcritical and supercritical aqueous solutions, more efforts should be made to develop a stable internal reference electrode which would reliably operate at temperatures above 300 °C. It seems that the development of an appropriate reference electrode is probably the most significant challenge if the electrochemical studies are to be performed in high-temperature subcritical and supercritical aqueous solutions.

Also, the development of a reliable and convenient pH sensor electrode that can operate reproducibly at temperatures above 300 °C also remains a goal of experimental electrochemistry. The desirable high-temperature pH sensor must provide a known, stable, and reproducible potential during the course of the measurement, which may involve long periods of time and harsh environments (oxidizing, reducing thermal shocks, etc.). Even though the recently developed flow-through YSZ(Hg/HgO) electrode can now be used at temperatures up to 400 °C, substantial effort is still needed to provide accurate laboratory-scale pH measurements to these extreme temperatures and to develop a rugged version of this electrode for industrial applications. In addition, reliable estimation of the diffusion and other undesirable potentials is a serious challenge in high-temperature electrochemical studies.

If one wants to obtain a comprehensive understanding of the interaction between a metal (or metal alloy) and a hydrothermal solution, then electrochemical kinetics and/or corrosion studies must be carried out. In particular, an electrochemical system capable of reliably operating at temperatures above 300 °C should be developed. It is a matter of fact that there are almost no data on the exchange current densities and the anodic and cathodic transfer coefficients for even the most fundamental electrochemical reaction in high-temperature subcritical and supercritical aqueous systems. Even the primary HERs and OERs have been poorly studied at temperatures above 100 °C. Therefore, the creation of a well-established method for measuring electrochemical kinetics and corrosion processes over a wide range

of temperatures remains a serious and unresolved experimental problem. Also, the proper control of the mass transport processes in electrochemical kinetics and corrosion experiments seems to be one of the main challenges of high-temperature electrochemical studies.

The buffer and reference solutions are necessary for high-quality potentiometric and electrochemical kinetics measurements, and it is highly desirable to estimate a set of the buffer solutions that can be used at temperatures above 100 °C. Thus far, little has been done for developing a necessary set of the high-temperature buffer systems.

In order to understand the surface chemistry at a solid oxide/water interface, high-temperature electrokinetic studies should be carried out. Recently, microelectrophoretic studies of the ZrO_2- and TiO_2-water suspensions to 200 °C showed that such measurements are possible, but substantial effort will be needed to extend these measurements to a temperature range above 300 °C. In addition, a possibility to study nanoparticulate aqueous suspensions at high temperatures would be very beneficial.

With regard to electric conductance measurements, significant progress has recently been achieved to accurately study aqueous solutions in high-temperature subcritical and supercritical conditions, and this method can now be considered a reliable approach for learning relatively simple chemical equilibria in the region of the critical point of water.

References

1. D. A. Palmer, R. Fernández-Prini, A. H. Harvey, (Eds.), *Aqueous Systems at Elevated Temperatures and Pressures: Physical Chemistry in Water, Steam and Hydrothermal Solutions*, Elsevier, Amsterdam, 2004.
2. S. N. Lvov, D. A. Palmer, in *The Physical and Chemical Properties of Aqueous Systems at Elevated Temperatures and Pressures: Water, Steam and Hydrothermal Solutions* (Eds.: D. A. Palmer, R. Fernández-Prini, A. H. Harvey), Elsevier, Amsterdam, 2004, p. 377.
3. S. D. Cramer, B. S. Covino, (Eds.), *ASM Handbook, Volume 13A, Corrosion: Fundamentals, Testing, and Protection*, ASM International, Materials Park, 2003.
4. G. G. Wildgoose, D. Giovanelli, N. S. Lawrence et al., *Electroanalysis* **2004**, *16*, 421.
5. D. Giovanelli, N. S. Lawrence, R. G. Compton, *Electroanalysis* **2004**, *16*, 789.
6. P. Kritzer, *J. Supercrit. Fluids* **2004**, *29*, 1.
7. A. J. Bard, L. R. Faulkner, *Electro-Chemical Methods. Fundamentals and Applications*, 2nd ed., Willey, New York, 2001.
8. D. R. Lide, Editor-in Chief, *Handbook of Chemistry and Physics*, 77th ed., CRS Press, Boca Raton, New York, London, Tokyo, 1996–1997.
9. G. B. Naumov, B. N. Ryzhenko, I. L. Khodakovsky, *Handbook of Thermodynamic Data*, US Geology Survey, Menlo Park, 1974.
10. M. H. Lietzke, R. W. Stoughton, *J. Phys. Chem.* **1963**, *67*, 2573.
11. D. D. Macdonald, *Corrosion* **1978**, *34*, 75.
12. S. N. Lvov, H. Gao, D. D. Macdonald, *J. Electroanal. Chem.* **1998**, *443*, 186.
13. S. N. Lvov, X. Y. Zhou, D. D. Macdonald, *J. Electroanal. Chem.* **1999**, *463*, 146.
14. D. A. Palmer, P. Bénézeth, D. J. Wesolowski, *Geochim. Cosmochim. Acta* **2001**, *65*, 2081.
15. K. E. Eklund, S. N. Lvov, D. D. Macdonald, *J. Electroanal. Chem.* **1997**, *437*, 99.
16. S. N. Lvov, X. Y. Zhou, G. C. Ulmer et al., *Chem. Geol.* **2003**, *198*(3–4), 141.
17. S. N. Lvov, X. Y. Zhou, M. V. Fedkin et al., *Geochim. Cosmochim. Acta* **2002**, *66*, A467.
18. X. Y. Zhou, S. N. Lvov, S. M. Ulyanov, United States Patent # 6,517,694, February 11, 2003.
19. S. N. Lvov, D. D. Macdonald, *J. Electroanal. Chem.* **1996**, *403*, 25.
20. S. N. Lvov, X. Y. Zhou, S. M. Ulyanov et al., *Chem. Geol.* **2000**, *167*, 105.
21. A. K. Covington, M. I. A. Ferra, Z. Y. Zou, *Electrochim. Acta* **1985**, *30*, 805.
22. P. Tremaine, K. Zhang, P. Benezeth, in *The Physical and Chemical Properties of Aqueous Systems at Elevated Temperatures and*

Pressures: Water, Steam and Hydrothermal Solutions (Eds.: D. A. Palmer, R. Fernández-Prini, A. H. Harvey), Elsevier, Amsterdam, 2004, p. 441.
23. D. J. Wesolowski, S. E. Ziemniak, L. M. Anovitz, in *The Physical and Chemical Properties of Aqueous Systems at Elevated Temperatures and Pressures: Water, Steam and Hydrothermal Solutions* (Eds.: D. A. Palmer, R. Fernández-Prini, A. H. Harvey), Elsevier, Amsterdam, 2004, p. 493.
24. K. Ding, W. E. Seyfried Jr., *Geochim. Cosmochim. Acta* **1995**, *59*, 4769.
25. K. Sue, K. Murata, Y. Matsuura et al., *Rev. Sci. Instrum.* **2001**, *72*, 4442.
26. K. Sue, K. Murata, Y. Matsuura et al., *Fluid Phase Equilib.* **2002**, *194–197*, 1097.
27. D. S. Seneviratne, V. G. Papangelakis, X. Y. Zhou et al., *Hydrometallurgy* **2003**, *68*, 131.
28. X. Y. Zhou, X. J. Wei, M. V. Fedkin et al., *Rev. Sci. Instrum.* **2003**, *74*, 2501.
29. M. F. Fedkin, X. Y. Zhou, J. D. Kubicki et al., *Langmuir* **2003**, *19*, 3797.
30. V. M. Tsionsky, L. B. Kriksunov, L. L. Krishtalik, *Electrochim. Acta* **1991**, *36*, 411.
31. L. N. Trevani, E. Calvo, H. R. Corti, *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 4319.
32. D. D. Macdonald, J. Mankowski, M. Karaminezhaad-Ranjbar et al., *Corrosion* **1988**, *44*, 186.
33. Z. Nagy, L. A. Curtiss, J. W. Halley et al., *J. Electrochem. Soc.* **1991**, *138*, 2032.
34. C.-Y. Liu, R. Shelly, A. J. Bard, *J. Phys. Chem. B* **1997**, *101*, 1180.
35. X. Y. Zhou, S. N. Lvov, X. J. Wei et al., *Corros. Sci.* **2001**, *44*(4), 841.
36. H. R. Corti, L. N. Trevani, A. Anderco, in *The Physical and Chemical Properties of Aqueous Systems at Elevated Temperatures and Pressures: Water, Steam and Hydrothermal Solutions* (Eds.: D. A. Palmer, R. Fernández-Prini, A. H. Harvey), Elsevier, Amsterdam, 2004, p. 321.
37. L. Hnedkovsky, R. H. Wood, V. N. Balashov, *J. Phys. Chem. B* **2005**, *109*, 9034–9046.

11
Electrochemistry of Silicon Etching

X. Gregory Zhang
Teck Cominco Metals Ltd., Mississauga, Canada

11.1 Introduction 751

11.2 Dissolution of Silicon 753
11.2.1 Fluoride Solutions 753
11.2.2 In Alkaline Solutions 755

11.3 Hydrogen Evolution 756

11.4 Dissolution Kinetics and Mechanisms 758

11.5 Reaction Mechanisms 760
11.5.1 A General Model 762
11.5.1.1 Elementary Steps 763
11.5.1.2 Reaction Paths 764

11.6 Etch Rates in Various Systems 767
11.6.1 General 767
11.6.2 In Fluoride-based Solutions 772
11.6.2.1 Absence of Oxidants 772
11.6.2.2 Effect of CrO_3 Addition 773
11.6.2.3 Effect of HNO_3 Addition 774
11.6.2.4 Addition of Other Oxidants 776
11.6.3 Alkaline Solutions 776
11.6.3.1 KOH Solutions 777
11.6.3.2 Reaction Kinetics 778
11.6.3.3 Other Inorganic Solutions 779
11.6.3.4 Hydrazine 780
11.6.4 Organic Solutions 780
11.6.4.1 EDP Solutions 780
11.6.4.2 Etching Reactions 781

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

11.6.4.3	Ethanolamine	782
11.6.4.4	Tetramethyl Ammonium Hydroxide (TMAH)	782
11.7	**Etch Rate Reduction of Heavily Doped Materials**	**783**
11.7.1	Mechanism	784
11.8	**Anisotropic Etching**	**785**
11.8.1	Rate-Limiting Process	787
11.8.2	Mechanism of Anisotropic Etching	788
11.8.2.1	A General Model for both HF and KOH Solutions	789
11.8.3	Basic Features of Anisotropically Etched Surfaces	792
11.9	**Surface Roughness**	**794**
11.9.1	Microroughness	795
11.9.2	Macroroughness	795
11.9.3	Crystallographic Characters and Formation of Hillocks	796
11.9.4	Origins of Roughness	798
11.10	**Applications of Etching**	**799**
11.10.1	Cleaning	799
11.10.1.1	RCA Cleaning	800
11.10.2	Defect Etching	800
11.10.3	Material Removal	801
11.10.3.1	Uniform Material Removal	801
11.10.3.2	Selective Material Removal	803
	References	804

11.1 Introduction

Etching of semiconductor crystal in solutions, that is, wet etching, is a dissolution process to uniformly or preferentially remove material from the crystal surface. The etching of silicon in solutions, with its large range of etch rates and great selectivity with respect to orientation, doping type and concentration, and foreign materials, has been widely used for micromachining, delineation of defects, surface cleaning and polishing, and so on, in silicon technology [1–3]. It will retain its importance in the future as silicon continues to be the mainstream material for new generations of electronic devices [4–6].

Numerous etching solutions and processes have been developed for silicon etching, which can be grouped into three major categories according to relative etching rates on surfaces of different crystallographic orientations and the uniformity of the etched surface: (1) isotropic etching system, represented by $HF-HNO_3$, (2) anisotropic etching system, represented by alkaline solutions, and (3) defect etching system, represented by $HF-CrO_3$ solutions as illustrated in Fig. 1. In each of these systems, a range of factors, such as illumination, potential, temperature, surface patterning, and so on, can be controlled to obtain the desired etch rate, surface condition, etched profile, and so on [3].

Wet etching is an electrochemical process in which charge-transfer process is involved across the crystal/solution interface. During etching, a reaction of oxidation of surface atoms of the crystal and a reaction of reduction of some chemical species in the solution concurrently occur. There can be three distinctive modes of the charge-transfer process during etching as illustrated in Fig. 2: (1) a net current passes through the global crystal/solution interface, (2) no net current passes through the global crystal/solution interface but currents flow within localized areas of the crystal and solution, (3) no current is involved anywhere in the crystal and solution beyond the scale of individual atoms. In the case of (1) the crystal has to be polarized by imposing a potential through an external circuit and the reduction reaction occurs on an electrode separated from the etching crystal. In the case of (2) and (3) no external polarization is required as the oxidation and reduction both occur at the open-circuit potential (OCP) on the same crystal surface. However, in case (2) the oxidation and reduction involve different surface atoms, while in case (3), both the oxidation and reduction reactions occur at the same surface atom.

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

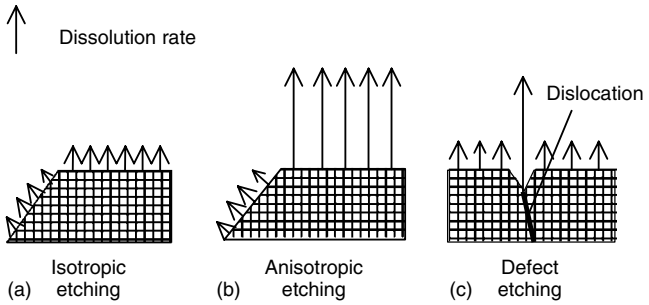


Fig. 1 Schematic illustration of the three major etching systems.

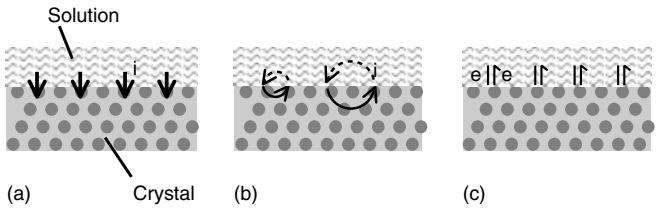


Fig. 2 Schematic illustration of the modes of charge-transfer process (a) polarized electrode with a net external current, (b) corrosion process with no net external current, and (c) chemical process with no local and external currents.

The processes in (1) and (2), involve a charge carrier on the surface, and depend on electrode potential, while that in (3) may not. Because of this difference, those in (1) and (2) are considered electrochemical reactions and that in (3) a chemical reaction. All these modes of charge transfer can be involved in an etching process but the relative contribution of each mode depends on the specific conditions. Identification of the charge-transfer mode and understanding of its relation to the structural and kinetic processes is a fundamental aspect of etching electrochemistry.

The etching systems for silicon etching have been extensively investigated since the 1950s when etching began to be used in device fabrication processes, accumulating an enormous amount of technical information. The details of this body of information are highly complex, considering the great diversity of circumstances under which the etching is performed, and the infinite possible combinations of solutions, materials, processing procedures, and structure/geometry factors which may be involved in a given circumstance.

However, the body of information generated from the studies on electrochemistry and that on etching are largely separated. This may be attributed to the different focuses and approaches in these two groups of studies; while the former focuses primarily on the scientific understanding of the electrochemical processes, the latter is largely motivated by providing engineering solutions to the industry. Also, at a different level, the information is segregated in terms of etching solutions for

HF and KOH, which are the two principal etching systems. There has been a lack of integration of the information on etching in these two solution systems.

This chapter aims to provide an overview of the experimental observations and theoretical developments involving silicon etching and to integrate the information on electrochemistry and etching as well as on KOH and HF. Also, it is an attempt to analyze the electrochemical processes involved in various etching phenomena from a global perspective. A conceptual approach is taken in dealing with the mechanistic aspects of the phenomena using the materials extracted from a comprehensive and detailed work on silicon electrochemistry published recently [3]. This chapter is organized roughly in two parts, with the first half relating to the electrochemistry and reaction mechanisms and the second half dealing with the specific etching phenomena.

11.2 Dissolution of Silicon

The dissolution of silicon can be best characterized by $i-V$ curves. Neglecting the details associated with silicon substrate such as doping, the current–potential relationship of silicon in aqueous solutions can be considered to be principally determined by the pH and HF concentrations as illustrated in Fig. 3. In nonalkaline and nonfluoride aqueous solutions, silicon as an electrode is essentially inert showing a very small current at anodic potential due to the presence of a thin oxide film. In alkaline solutions silicon is also passivated by an oxide film at anodic potentials but is active below the passivation potential, V_p. In fluoride solutions silicon electrode is active in the entire anodic region as shown by the large anodic current.

11.2.1 Fluoride Solutions

The $i-V$ curves in HF solutions are different for p-Si and n-Si in the dark due to the difference in the concentrations of holes on the surface of these two materials, as shown in Fig. 4 [7]. Large currents can be obtained on p-Si by anodic polarization to increase the concentration of holes at the surface. On the other hand, for nondegenerated n-Si the anodic current is limited by the availability of holes. The $i-V$ curves for n-Si become identical to that of p-Si when n-Si is illuminated at a high light intensity.

The anodic polarization curves of p-Si or strongly illuminated n-Si in fluoride solutions are typically characterized by two peak currents, J_1 and J_3, and two plateau currents, J_2 and J_4, as shown in

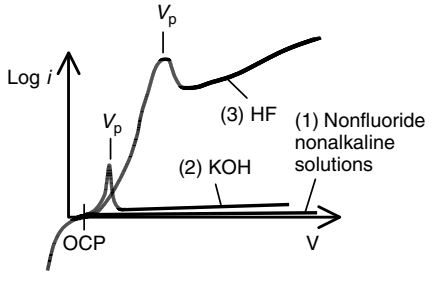


Fig. 3 Typical $i-V$ curve in HF and nonfluoride, nonalkaline solutions.

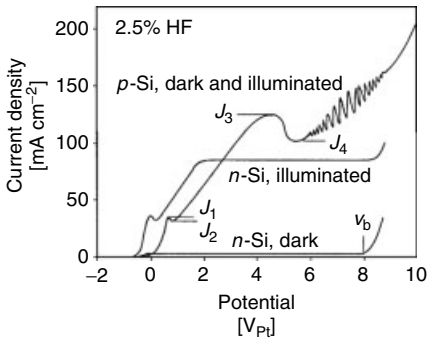


Fig. 4 Current–potential curves for dark and illuminated *p*- and *n*-Si electrodes in a 2.5 wt% HF electrolyte (after Lehmann and Foll [7]).

Fig. 4 [7, 8]. At anodic potentials near J_1 the electrode behavior is characterized by an exponential dependence of current on potential and by the uneven dissolution of silicon surface leading to the formation of porous silicon (PS) [9]. The values of the characteristic currents J_1 to J_4 are a function of electrolyte composition but are largely independent of doping. At potentials more positive than the second plateau current J_4, current oscillation may occur [8]. The surface resulting from dissolution at potentials higher than the second peak is brightly smooth, while that produced between the first and second peak is relatively less smooth [10].

Figure 5 shows that the current increases exponentially with increasing potential from OCP [9]. It breaks off from the exponential behavior at larger overpotentials, exhibits a peak, J_1, and then attains a relatively constant value at J_2. Examination of the surface anodized at different potentials, indicates that formation of PS occurs in the exponential region but not at potentials more positive than the peak potential. The potential corresponding to the maximum slope of the i–V curve is about the upper limit for formation of a uniform PS layer. At potentials between the maximum slope and the current peak, a porous layer may still form, but its surface coverage is not uniform.

Visible hydrogen evolution occurs in HF solutions at anodic potentials in the exponential region anodic of OCP [9, 11]. The rate of hydrogen evolution substantially decreases as potential approaches the current

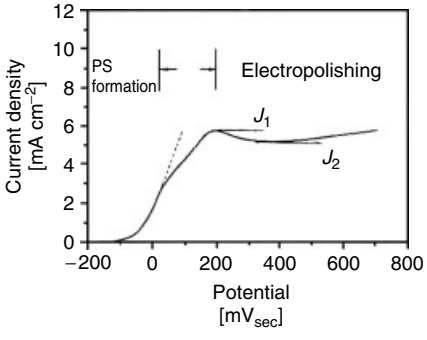


Fig. 5 Current–potential curve of p^+ silicon sample in 1% HF solution with potential sweep rate of 2 mV s^{-1} (after Zhang et al. [9]).

peak, J_1. Hydrogen evolution ceases above the current peak.

The Tafel slope determined in the exponential region in HF solutions is about 60 mV/decade for *p*-Si and n^+-Si. For lowly doped *n*-Si, since illumination is required to generate electrode reactions, the exponential region occurs only at sufficiently high illumination intensities when the generation of photoholes is no longer the rate-limiting step. Under a highly illuminated condition, the Tafel slope of *n*-Si is similar to that of other doping conditions. The 60-mV Tafel slope may be due to either the transport of charge carriers in the space charge layer or charge transfer in the Helmholtz double layer [3, 9]. The anodic behavior is identical for *p*- and *n*-type materials when the rate of carrier transfer between the surface and the semiconductor bulk is sufficiently fast implying that the reaction processes are determined by the chemical nature of silicon material independent of the electronic character of silicon.

For *n*-Si in the dark, the current at anodic potentials is characteristically small due to the reverse bias. The onset potential at which current rises sharply in the dark marks the breakdown of the barrier layer and initiation of interface tunneling on the silicon surface [12]. The breakdown potential, V_b (shown in Fig. 4), depends on the doping level of the material: the lower the doping the higher the breakdown potential [10, 13–15]. At the potentials negative of the breakdown potential, the anodic reactions require illumination to generate holes.

The effective dissolution valence of silicon, *n*, that is, the average number of electrons flowing through the external circuit per dissolved silicon atom, has been found to vary with silicon material, solution composition, anodic polarization, and illumination conditions [3]. In fluoride containing solutions, the dissolution valence varies from 2 in the exponential region to near 4 in the electropolishing region. For *n*-Si, the dependence of dissolution valence on potential varies with illumination intensity. At high illumination intensity, when the process is not light intensity limited, the variation of dissolution valence with potential is observed to be similar to that of *p*-Si. The dissolution valence may be lower than 2 at low current densities due to chemical dissolution of PS. The value in the electropolishing region is somewhat lower than 4 due to the chemical oxidation of the silicon reaction intermediates.

Since hydrogen evolution occurs at anodic potentials and it is responsible for an apparent dissolution valence smaller than 4, the effective dissolution valence can be used as a measure for the hydrogen efficiency. A silicon dissolution valence of 2 can be used to indicate 100% efficiency for hydrogen evolution, that is, for every dissolved silicon atom one hydrogen molecule is generated. Figure 6 shows the effective dissolution valence and hydrogen evolution in HF solutions under different conditions.

11.2.2 In Alkaline Solutions

The anodic behavior of silicon electrodes in alkaline solutions is characterized by the formation of an oxide film and passivation of the surface. The typical *i*–V curves of *n* and *p* types of silicon in KOH solution are shown in Fig. 7 [16]. Silicon surface is passivated at an overpotential marked by a small current peak. The potential at the current peak is called the *passivation*

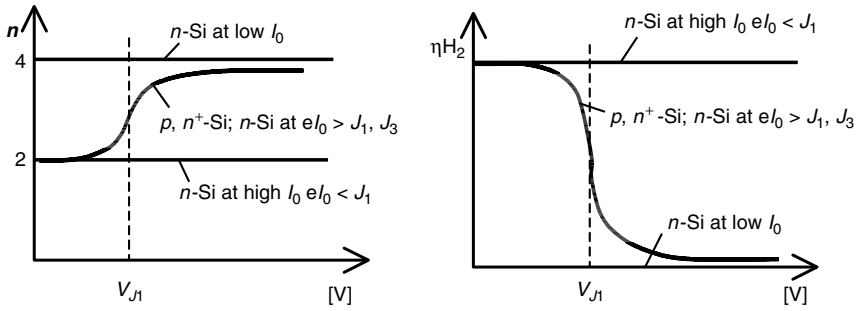


Fig. 6 Summary of effective dissolution valence and hydrogen evolution efficiency in HF solutions as a function of potential for different materials and illumination conditions, V_{J1} – potential of first current peak J_1; I_0 – light intensity.

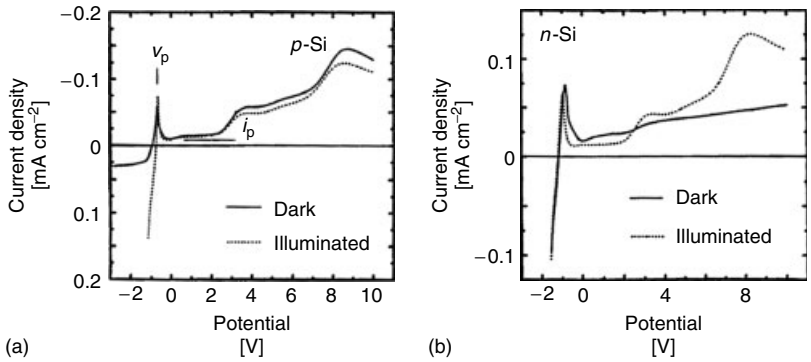


Fig. 7 *i*–V curves in 2 M KOH in the dark and illumination, (a) *p*-Si, (b) *n*-Si (after Chen et al. [16]).

potential, V_p, and in the systems shown in Fig. 7 it is −0.92 and −0.71 V_{sce} for *n*- and *p*-type silicon respectively. The small current at the potentials more positive than V_p is the passivation current, i_p which reflects the dissolution rate of silicon in the passive region. Also, the *i*–V curves are essentially identical for samples with different doping levels [17]. The data show, similar to the fluoride solutions, that the basic features of the anodic behavior in alkaline solutions are determined by the chemical nature of the silicon material but not by the electronic properties.

In alkaline solutions, the effective dissolution valence at OCP is zero and changes only very slightly with anodic polarization before the passivation peak; it is less than 0.4 at V_p [18]. The dissolution reaction below the passivation potential is thus almost completely chemical in nature [18–20].

11.3 Hydrogen Evolution

In aqueous solutions at OCP in the absence of redox couples, hydrogen evolution is

the only reduction reaction accompanying the dissolution of silicon. In HF solutions, H_2 evolution at cathodic potentials in the charge-transfer process is found to proceed almost entirely over the conduction band [7–11]. The current increases exponentially with decreasing potential. The exponential i–V relationship exists also on illuminated p-Si at currents that are significantly lower than the saturation value as shown in Fig. 8 [21]. The Tafel slope for hydrogen reaction for all types of silicon materials in the dark or under illumination appears to be between 140 and 200 mV/decade [22]. This is much higher than the 60 mV/decade required for the processes that are limited by the supply of electrons from the semiconductor, which suggests that the charge-transfer reaction involved in the hydrogen evolution at cathodic potentials is mainly controlled by the processes in the Helmholtz layer. The exchange current density for hydrogen reduction is in the range of 10^{-6}–10^{-11} A cm^{-2} depending on the specific conditions of silicon/electrolyte system [22–24].

The reaction kinetics of hydrogen evolution is particularly sensitive to trace amounts of noble metal impurities in the solution such as Cu and Au, which tend to deposit on the silicon surface [24, 25]. Figure 9 shows the effect of different metal deposits on the photocurrent of p-Si in a H_2SO_4 solution [26]. There appears to be a correlation between the effect of these metals on photocurrent and the exchange current density of hydrogen evolution on the metals. In addition to metal deposits, hydrogen evolution can also be catalyzed through deposition of a layer of metal

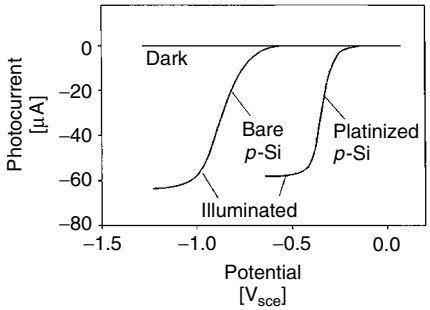


Fig. 8 Steady-state photocurrent–voltage curves for naked p-type Si and for platinized p-Si. Aqueous solution buffered to pH 6.6; illumination is at 632.8 nm, 2.5 mW cm^{-2}. The platinized p-type Si was prepared by photoelectrochemical redution from 1×10^{-3} M KPtCl in 0.1 M NaClO/H_2O at −0.3 V_{sce} until 1.1×10^{-2} C cm^{-2} had passed (after Dominey et al. [21]).

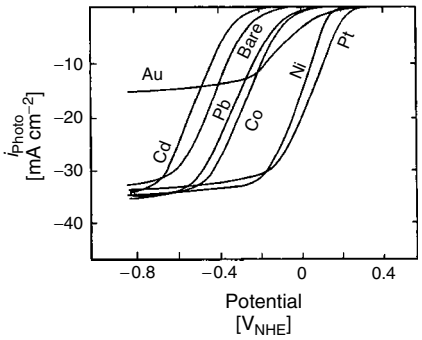


Fig. 9 Cathodic i–V curves on p-Si deposited with different metals in 0.5 M H_2SO_4 (after Szklarczyk and Bockris [26]).

or polymer on the silicon surface or by addition of a redox couple in the solution [27–29].

Hydrogen evolution on silicon may proceed chemically or electrochemically depending on the conditions. Hydrogen evolution near OCP and at anodic potentials can proceed completely chemically, that is, without involving the carriers from the electrode. The chemical nature of hydrogen evolution is responsible for less than 4 of the silicon effective dissolution valence as shown in Fig. 6. A change from a chemical process to an electrochemical process occurs when the potential varies from anodic values to cathodic values as schematically illustrated in Fig. 10. Hydrogen evolution at cathodic potentials is predominantly electrochemical due to the lack of silicon dissolution and abundance of electrons on the surface on *n*-Si or illuminated *p*-Si.

The efficiency of hydrogen evolution at anodic potentials is near 100% (ratio of H_2:Si equals 1) in the exponential region but drops to about 50% at the peak potential [30, 31]. At potential larger that the current peak, a considerable number of hydrogen molecules, corresponding to 10 to 13% efficiency, are generated [10–15, 32]. In this region, silicon dissolution proceeds with oxide formation and dissolution of the oxide. Hydrogen evolution means that some silicon atoms are only partially oxidized in the oxide-forming stage and are fully oxidized by hydrogen afterwards.

In KOH, at OCP and at anodic potentials lower than V_p, there is no participation of charge carriers and the reactions, hydrogen evolution, and dissolution of silicon, are almost 100% chemical such that the etching of silicon is characterized with the dissolution of one silicon atom and the evolution of two hydrogen molecules. At potentials higher than V_p the surface is passivated and both silicon dissolution and hydrogen evolution cease [20]. At cathodic potentials, hydrogen evolution on *p*-Si is also chemical due to the lack of electrons. However, for *n*-Si at cathodic potentials hydrogen evolution is mainly electrochemical due to the abundance of electrons from the semiconductor.

In the presence of redox couples, the relative contribution and mechanism of hydrogen evolution may change, the detail of which can be found in Ref. 3.

11.4 Dissolution Kinetics and Mechanisms

The conduction of the charge carriers in the space-change layer during anodization of silicon can be due to either electron injection into the conduction band or hole capturing from the valence band. For *p*-type silicon, the anodic reactions in HF and KOH solutions proceed predominantly over the valence band [7, 11,

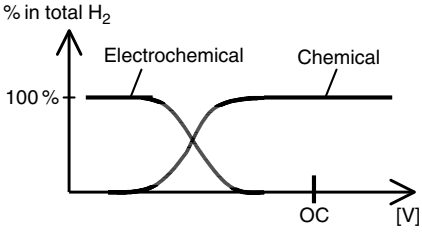
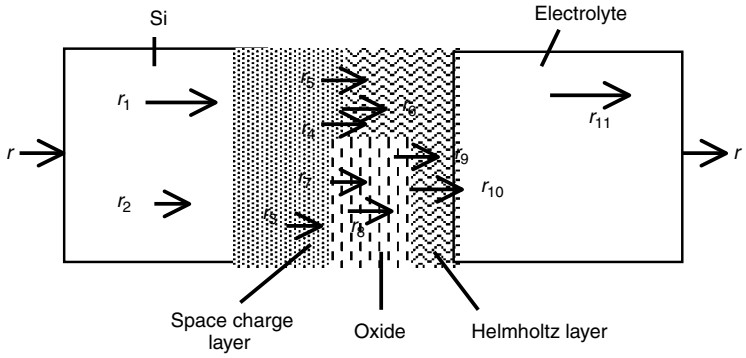


Fig. 10 Schematic illustration of the relative contributions of chemical path versus electrochemical path in H_2 evolution (after Zhang [3]).

18, 20, 33]. For *n*-Si, anodic reactions in the dark at current densities higher than the limiting current is a conduction band process but can involve both the conduction band and the valence band under illumination. The relative participation of conduction band versus valence band is a function of light intensity; the conduction band dominates at a low light intensity and the valence band at a high light intensity [34–36]. Participation of holes and electrons in the dissolution reactions varies with the type of silicon, illumination conditions, and electrode potential.

The mechanism of charge transfer across the interface barrier layer is different for lowly doped and heavily doped *p*-type silicon. For lowly *p*-type doped silicon the process is by thermal emission of holes to go over the barrier layer while it is by Zener tunneling for heavily doped materials [11, 37]. For *n*-Si the conduction band processes depend on doping density and on illumination intensity. For heavily doped *n*-Si it is by Zener tunneling and the *i*–V curve is identical to that of *p*-Si. For moderately or lowly doped *n*-Si in the dark the reaction is limited by the minority holes, which are required to initiate the dissolution process. Significant dissolution of *n*-Si can proceed when a large amount of holes are generated by illumination. Large current can also be generated on *n*-Si in HF solutions in the dark at relatively high anodic potentials due to an interface tunneling process [10–12].

Each of the possible processes, as schematically illustrated in Fig. 11, in the multilayer silicon/electrolyte interface region can be the rate-limiting process under certain conditions. For example, as has been summarized in Ref. 3 the anodic reaction processes on *n*-Si in the dark are limited by the minority hole transport in the bulk of silicon, that is r_2. For *p*-Si and illuminated *n*-Si in fluoride solutions at potentials negative of the first current peak, J_1, the reaction



r_1 and r_2 – majority and minority carrier transport, r_3 – transport of holes to the surface
r_4 – charge transfer across the Helmhotz layer, r_5 – electron injection
r_6 – chemical dissolution, r_7 – oxide formation, r_8 – ionic transport in oxide,
r_9 – injection of oxidants, r_{10} – dissolution of oxide,
r_{11} – mass transport in electrolyte

Fig. 11 Schematic illustration of the processes involving the transport of charge and species in the different phases in the Si/electrolyte interface region (after Zhang [3]).

Tab. 1 Rate-limiting steps involved in the anodic oxidation of silicon[d]

	HF				KOH			
	r[a]	p	n	p^+, n^+	r[a]	p	n	p^+, n^+
Dark								
OCP	~0	r_3	r_3	r_3	Large	r_6	r_6	r_6
OCP $< V < V_p$	Large	r_4	r_2 or r_5[b]	r_4	Large	r_6	r_6	r_6
$V \geq V_p$	Large	r_{10}, r_{11}	r_{10}, r_{11}	r_{10}, r_{11}	Small	r_{10}	r_{10}	r_{10}
Illuminated								
OCP	Large	r_4	r_3	r_4	Large	r_6	r_6	r_6
OCP $< V < V_p$	Large	r_4	r_3 or r_4 or r_5[c]	r_4	Large	r_6	r_6	r_6
$V \geq V_p$	Large	r_{10}, r_{11}	r_{10}, r_{11}	r_{10}, r_{11}	Small	r_{10}	r_{10}	r_{10}

[a] Oxidation rate, qualitative comparison to the dark limiting current density.
[b] It is r_2 below the breakdown potential but it is r_4 above the breakdown potential.
[c] It is r_3 when the current equals the saturation photocurrent, is r_5 when the current is larger than the saturation photocurrent and is r_4 when the current is less than the saturation photocurrent.
[d] In aqueous solutions containing no redox couples other than the ones associated with HF, KOH, and H_2O; presence of redox couples can affect the rate-limiting process [3].

rate is determined by the charge-transfer process across the electrode/electrolyte interface, that is, r_4 and r_5. At potentials positive of J_1, which is the electropolishing region, the rate-determining step in the anodic reaction is the dissolution of the anodic oxide film, that is r_{10}. The dissolution of the oxide film formed at low fluoride concentrations is mainly kinetically controlled, that is r_{10}, while for high fluoride concentrations the process is mainly diffusion controlled, r_{11}. There is a critical concentration, depending on pH, rotation rate, and potential, at which the contributions by kinetic and diffusion processes are equal [8].

In KOH solutions, the rate-limiting process at OCP is of chemical nature, that is, only r_6 is involved. The electrochemical processes r_4 and r_5 are increasingly involved as potential is increased from OCP to V_p. At potential larger than V_p the anodic reaction is limited by the dissolution rate of the oxide r_{10}, which, unlike in fluoride solutions, is completely a surface-controlled process in KOH solutions. Table 1 summarizes the rate-determining processes identified for the anodic reactions on silicon under different conditions in KOH and HF.

11.5 Reaction Mechanisms

The reactions involved in the anodization of silicon in aqueous solution, depending on whether fluoride ions are present, as originally proposed by Turner [38], and Memming and Schwandt [11] in the 1950–1960s, have two different paths determined by the two essential reactants, H_2O and HF. Later, Matsumura and Morrison [34, 39] observed current doubling and quadrupling on illuminated *n*-Si in HF solutions and suggested that during oxidation of Si to Si^{4+} all the steps, except perhaps for the first, occur via electron injection into the conduction band.

Whether current doubling and quadrupling occur is determined by the competing processes depending on the availability of holes. The observation of hydrogen termination of silicon surface in HF solutions in the 1980s led to the conclusion that the dissolution of silicon starts by adsorption of hydrogen and then the adsorbed hydrogen atoms are replaced by fluoride atoms. The fluoride terminated silicon atoms further react with HF molecules to break the silicon atoms from the lattice and dissolve into the solution [32, 40–42]. The reaction of silicon atoms with hydrogen results in hydrogen evolution, which is responsible for the chemical component of the dissolution reaction and for an effective dissolution valence less than 4. The relative contribution of chemical reactions and electrochemical reactions depends on pH of the fluoride solution [41].

The individual steps in the dissolution process in HF solutions were further considered in a number of studies in the late-1980s and early-1990s to involve a hole capturing in the first step followed by an electron injection as the second step [30, 42, 43]. The oxidized silicon intermediate, having a valence of two, is further oxidized by the injection of two electrons or reacting with the hydrogen ions depending on light intensity. More recently, an intermediate Si(I), which is an electron deficient surface silicon atom and is mobile on the surface, was proposed to act as a radical to catalyze the hydrogen reduction reaction and to account for the transition of the dissolution reaction from valence 4 to valence 2 with increasing light intensity [44].

In alkaline solutions, an important issue involving the reaction mechanisms has been whether the etching process at OCP is of chemical (Fig. 2c) or electrochemical nature (Fig. 2b), that is, whether charge-transfer processes involve carriers in the electrode. Some studies [45, 46] suggested that the dissolution in KOH at OCP is an electrochemical process with concurrent anodic dissolution of silicon and reduction of hydrogen ions. The oxidation of silicon gives out electrons that are consumed for the reduction of hydrogen. Both OH^- and H_2O are the active species in which OH^- is involved in silicon dissolution and H_2O in hydrogen evolution. On the other hand, other people [47–49] have proposed that silicon etching in KOH solutions at OCP is of chemical nature and the etching reactions involve H_2O attack on Si–Si bonds to form Si–H and Si–OH bonds followed by OH^- attack to form $Si-O^-$ and H_2.

The information on electrochemistry, as summarized in the previous sections, appears to support the chemical mechanism for the etching in KOH at OCP. It changes, however, with potential; electrochemical reaction starts to participate and increases with increasing potential up to the passivation potential, above which it is almost completely an electrochemical reaction [19]. Also, the silicon surface in KOH solutions is found to be terminated with hydrogen similar to that in HF solutions, which means that the first step must then involve a dissociation of the Si–H bond by hydrolysis to form a Si–OH bond [19, 50].

A systematic review of the historic development of the theories on the mechanisms indicates that although numerous models on the dissolution of silicon have been proposed for etching in both HF and KOH solutions, each group of models is largely detached from the other. Yet, a complete model must be able to explain coherently the continuum of the dissolution process varying from that in acidic HF solutions to that in alkaline KOH solutions. The following analysis is an attempt for such a model integrating the previous findings

in the electrochemistry of silicon in both KOH and HF.

11.5.1 A General Model

Table 2 is a summary of the characteristic features of the reactions occurring on silicon in HF and KOH, in terms of nature of reaction, rate, effective dissolution valence, photoeffect, and uniformity of the surface.

Although there are many differences in HF and KOH solutions as shown in Table 2, the overall reactions in the two systems are similar in two important aspects:

Tab. 2 Characteristics of silicon electrode in HF and KOH solutions at different potentials [3]

	HF	***KOH***
Dissociation of solutes	$K = 7–13 \times 10^{-3}$	Complete
Concentration of OH^-	Extremely low	High
Activity of H_2O	High	Low at high [KOH]
Solubility of SiO_2	High	High
Rate of SiO_2 dissolution	High	Low
Hydrogen termination at OCP	Yes	Yes
Dissolution rate of Si at OCP	None	High
$OCP \leq V < V_p$	Increase with V, lowest at OCP	Decrease with V, highest near OCP
at $V > V_p$, r_p	High, $r_p \gg r_{ocp}$	Low, $r_p \ll r_{ocp}$
Effective dissolution valence		
at OCP	No dissolution	~0
$OCP < V \leq V_p$	2–4	<1
low light intensity, *n*-Si	~4	<1
high light intensity, *n*-Si	~2	<1
at $V > V_p$	~3.6	~4
Chemical reaction	H_2 evolution	H_2 evolution
at OCP	Low	High, ~100%
$OCP < V \leq V_p$	High, <50%	High, >50%
at $V > V_p$	~10%	~0
Photocurrent quantum efficiency, *n*-Si		
$OCP \leq V \leq V_p$		
low light intensity	~4	~0
high light intensity	~2	~0
$V > V_p$	1	1
Difference in *n*-Si and *p*-Si near OCP	Yes	No
Band bending at OCP	*p*-Si downward, *n*-Si upward	*p*-Si and *n*-Si downward
Dependence on orientation		
at $V < V_p$	Yes	Yes
at $V > V_p$	No	No
Dissolution uniformity		
at $V < V_p$	Porous silicon formation	Anisotropic
at $V > V_p$	Uniform	Uniform

r_p: dissolution rate in passive region; r_{ocp}: dissolution rate at OCP.

the silicon surface is dynamically terminated by hydrogen and breaking of the silicon–silicon backbond is facilitated by the adsorption of the electronegative ligands such as F^- or OH^-. More specifically this means: (1) the initial surface is hydrogen terminated, (2) the Si–Si backbond requires that the hydrogen termination is first replaced by F^- or OH^-, (3) the silicon atoms on the newly exposed layer are terminated by hydrogen so that the surface after the dissolution of one silicon layer is identical to that before the dissolution.

A key aspect of the reaction scheme is the valence state of the silicon and hydrogen atoms at different stages of the various reaction paths. The nonpolarized nature of the Si–H bond indicates that the hydrogen atom bonded to the silicon atoms on the surface is at a reduced state with a valence of 0. The lattice silicon atoms bonded to hydrogen atoms also have a valence of 0 (or four valence electrons). The silicon atoms on the surface that are bonded to one hydroxyl ion or one fluoride ion, on the other hand, have a valence state of +1.

11.5.1.1 Elementary Steps

The dissolution of a surface silicon atom involves the replacement of the surface hydrogen atom by an OH^- in KOH solutions and by F^- in HF solutions. The difference between the hydrogen replacement by OH^- and that by F^- is whether holes are involved. As shown in Fig. 12, in KOH, upon the hydrogen replacement by OH^-, the silicon atom becomes a radical, which in the following reaction steps is neutralized by the reduction of hydrogen ions. In HF, the replacement of hydrogen by F^- requires a hole, which results in a neutralized Si–F bonding. The valence state of the adsorbed hydrogen before and after the replacement, being zero, is not changed in both solutions. The important difference is that after being terminated by an OH^- ion, the silicon atom is a radical with two extra electrons, while in Si–F termination the bonds associated with the silicon atom are neutralized.

The hydrogen adsorption onto a silicon atom is a reduction process since the valence of the hydrogen atom is changed from +1 to 0. It occurs when the backbond of Si–SiF or Si–SiOH is broken by

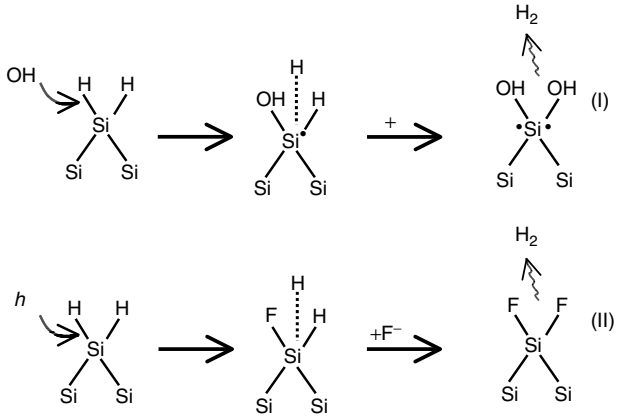


Fig. 12 Elementary steps in the replacement of H by OH^- or F^-.

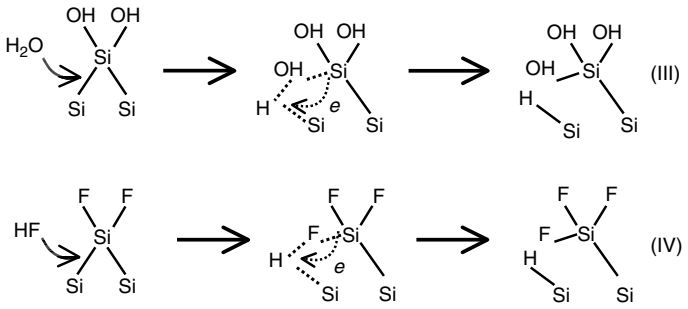


Fig. 13 Elementary steps involved in the breaking of Si—Si backbond and termination by hydrogen.

reacting with H_2O or HF. Transfer of one electron from the Si—OH bond to the hydrogen of the Si—H bond must then occur according to the process as illustrated in Fig. 13. An important feature of this process is that no carriers from the solid are involved and it is therefore a chemical reaction in nature. This is the key reaction step responsible for the chemical character of the dissolution process in which a hydrogen ion is reduced by an electron from the Si—Si bond. The breaking of the Si—Si backbond by any group other than HF and H_2O does not preserve the local charge neutrality (the two specific atoms involved in the bonding) and has to involve carriers from the bulk. This could explain why hydrogen adsorption onto silicon does not occur easily by the dissociated hydrogen ions in the solution but rather by H_2O or HF molecules through the attacking of the Si—Si backbonds.

Therefore, OH^- and H_2O are the reactants involved in KOH solutions and F^-, HF, and H_2O are involved in HF solutions. OH^- and F^- are responsible for initiating the attack by replacing the adsorbed hydrogen atoms while H_2O and HF are responsible for attacking the Si—Si backbonds. The elementary steps involving these reactions shown in Figs 12 and 13 can account for the essential features, that is, dynamic hydrogen termination of the silicon surface and breaking of the silicon backbond due to adsorption of fluoride or hydroxyl ions. The relative contributions of these processes in the reactions can then account for the variations in the effective dissolution valence and in the quantum efficiency under different conditions. An important aspect is that electronic carriers in the silicon semiconductor do not affect the chemical nature of the reactions so that the reactions shown above are the same on *p*- and *n* type of materials. The electronic carriers, however, affect the rate and the path of the reactions, as will be discussed in the following section.

11.5.1.2 **Reaction Paths**

Figures 14 and 15 show the possible reaction schemes for the various situations involving silicon electrodes in HF and KOH solutions. These different paths can account for the many details experimentally observed in the dissolution or passivation of silicon in HF and KOH. Table 3 is a summary of the reaction paths involved in different potential ranges in HF and KOH solutions. Variation of the reaction paths

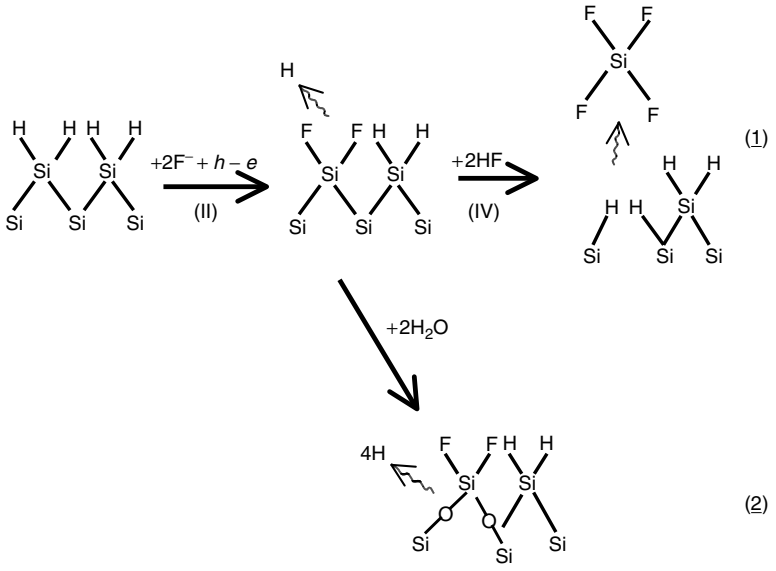


Fig. 14 Possible reaction paths in HF solutions at anodic potentials.

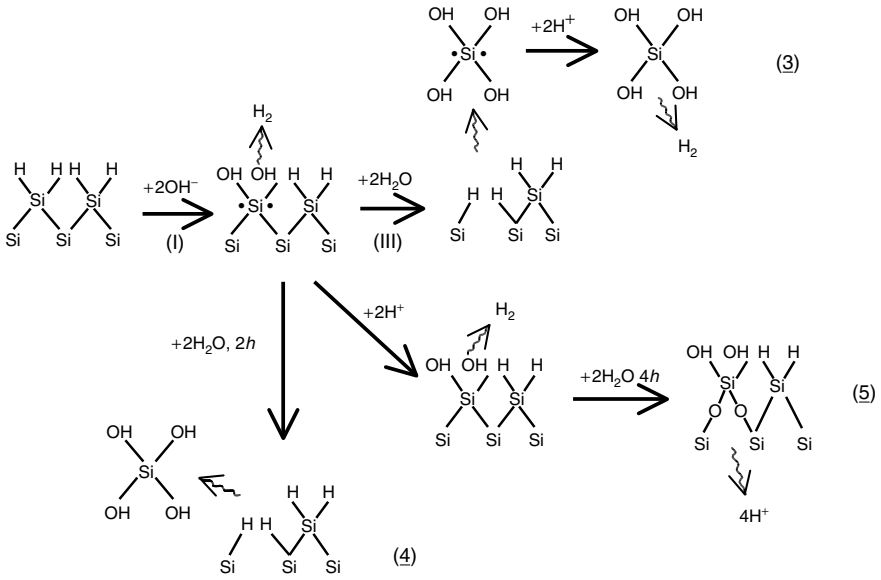


Fig. 15 Possible reaction paths in KOH solutions.

Tab. 3 Effective dissolution valence and conditions for different reaction paths shown in Figs 14 and 15

Path	*n*	*Conditions*
		HF
(1)	2	OCP $< V \ll V_p$
(1) and (2)	$2 < n < 4$	OCP $\ll V < V_p$
(2)	4	$V \geq V_p$
		KOH
(3)	0	OCP
(3), (4), (5)	<1	OCP $< V < V_p$
(5)	4	$V > V_p$

n: effective dissolution valence.

from those shown in Figs 14–15 are also possible based on the elementary steps outlined in the previous section.

Reaction paths (1) and (2) in Fig. 14 account for the anodic reactions on *p*-Si and illuminated *n*-Si in HF solutions at high light intensities. Path (1) is involved in the exponential region at an anodic potential much lower than V_p responsible for direct dissolution of silicon and dissolution valence of 2, while path (2) is involved at a potential above V_p responsible for the indirect dissolution of silicon through formation and dissolution of oxide and for the dissolution valence of 4. At a potential that is lower than but close to V_p both paths (1) and (2) are involved resulting in a dissolution valence between 2 and 4. There is also a possible reaction, parallel to (1) and (2), involving surface defects, which occurs near OCP and is responsible for the dissolution valence and quantum efficiency of 4 observed on *n*-Si at low light intensities [3].

Figure 15 shows the reaction paths in KOH solutions. Path (3), which involves no holes and electrons, is responsible for the chemical dissolution in KOH solutions at OCP. Path (4) involves two carriers and is responsible for the electrochemical part of the dissolution reaction at anodic potentials below V_p. Path (5) results in the formation of Si−O−Si bond and is responsible for the oxide formation and passivation at potentials above V_p.

Except for reaction path (3), which is purely chemical in nature, all the other reaction paths are of electrochemical nature, at least partially. These electrochemical reactions depend on the carrier transfer between the states at the interface and those in the semiconductor and thus their rates increase with increasing potential or illumination. While the reaction paths (1), (3), and (4) result in the direct dissolution of silicon, the reaction paths (2) and (5) result in the formation of Si−O−Si bonds, the dissolution of which results in an indirect dissolution path. The rate of reaction paths (2) and (5) also increase oxide formation with potential. As the coverage of the surface by Si−O−Si bonds increases with increasing potential, the surface becomes increasingly less active and becomes passivated when these bonds fully cover the surface. Further reaction has to proceed via the breaking of Si−O−Si bonds, which is fast in HF solutions but very slow in KOH solutions.

The surface of the oxidized silicon layer is terminated by OH in KOH but may be terminated by both OH and F in HF. In HF on the surface covered by an anodic oxide layer, adsorption of OH is required for the growth of the oxide, while adsorption of F is required for the dissolution of oxide. The Si−O−Si bonds are rather stable in KOH such that the dissolution rate in the passive region is very low. On the other hand, the Si−O−Si bonds are not stable in HF

due to the attack by the fluoride species and the dissolution rate is high at the potential larger than that of the current peak.

For the anodic dissolution of silicon in HF solutions, the first step involves replacing hydrogen with fluoride and the rate is proportional to the product of $[H_{Si}] \times [F^-_{solution}] \times [hole_{surface}]$ according to reaction (1) where H_{Si} is adsorbed hydrogen. The rate of the following step, involving breaking the Si−Si−F backbond by HF, is proportional to [Si−F] × [HF]. On the other hand, for the chemical dissolution of silicon in KOH solutions, the rate of hydrogen replacement is proportional to $[H_{Si}] \times [OH^-_{solution}]$ without involving holes and the breaking of the Si−Si−OH backbond is proportional to [Si−OH] × $[H_2O]$ according to reaction (3). A major difference between the reactions in the two solutions is that the hydrogen-replacing step in HF requires holes while in KOH it does not. Thus, it can be seen that in HF solutions the reaction rate is low at OCP but rapidly increases with potential. On the other hand, in KOH solutions the rate of the dissolution reaction is high but does not increase with potential. Because of the high concentration of OH^- in KOH, the rate of reaction (3) is high, but is very low in HF solutions because of the extremely low OH^- concentration. This explains the low dissolution rate in HF and high rate in KOH at OCP.

According to reactions (1) and (3), the concentration of hydrogen ions in both solutions does not play a direct role in the reaction processes. This is because the generation of adsorbed hydrogen, H_{Si}, does not depend on the concentration of hydrogen ions but on [HF] in HF solutions or $[H_2O]$ in KOH solutions. This explains why high dissolution rates are observed in KOH solutions, in which the concentration of hydrogen ions is extremely low.

11.6 Etch Rates in Various Systems

11.6.1 General

The two principal etching solution systems for silicon are HF solutions and alkaline solutions. This is because silicon is inert in aqueous solutions due to the formation of an insoluble surface oxide except for HF solutions or alkaline solutions in which the oxide is soluble. Various chemical agents can be added into these two solutions to obtain a control on etch rate, etch selectivity, solution stability, and quality of the etched surface. One major difference between these two systems is that the etch rate of silicon HF solutions is similar among the various crystalline orientations, that is, isotropic, while in alkaline solutions it strongly depends on the crystalline orientation, that is, anisotropic. Another difference is that silicon oxide, which may be present on the silicon surface prior to or during an etching process, etches fast in HF solutions while it etches very slowly in alkaline solutions relative to the etch rate of silicon.

Examples of the specific etch rates in various solutions reported in the literature are given in Table 4. Several points may be made regarding the data in this table: (1) silicon can etch at a wide range of rates, as much as 9 orders of magnitude, (2) the highest etch rates are observed in HF solutions, (3) the etch rates of the three major crystal planes vary only marginally in HF solutions but vary greatly in alkaline solutions, (4) etch rate is specific to a given set of etching conditions. In addition

Tab. 4 Etch rates, A s^{-1}, of single silicon crystals silicon in various solutions

Si[a]	*Solution*[c]	*T*	*(100)*	*(111)*	*(110)*	*References*
	HF solutions					
n, 2 Ω	9.6% HF, 15 hours	rt	–	8×10^{-3}	–	51
p, 1 Ω	48% HF, 5 days	rt	4×10^{-4}	3×10^{-4}	–	52
n, 3 Ω	1 M NH_4F, pH = 4.2	rt	–	0.03	–	53
	1 M NH_4F, pH 8.25	rt	–	10^{-3}	–	54
	HF + oxidants					
Poly Si	0.2% HF + 50% HNO_3	25	0.83	–	–	55
Poly Si, 10^{-10}	2% HF + 50% HNO_3	rt	115	–	–	56
p	1.3% NH_4F + 5% H_2O_2	78	–	0.3	–	57
p	0.1% HF + 10% H_2O_2	rt	0.05	–	–	58
n, 3 Ω	33.5% HF + 23.1% HNO_3	rt	280 000	–	–	59
n, 3 Ω	10% HF + 56% HNO_3	rt	14 000	13 300	12 700	60
n, 0.002, 3 Ω	1% HF + 3.8 $\times 10^{-4}$M $KMnO_4$	rt	70	–	–	61
9–23 Ω	0.01M CrO_3 + 2M HF	rt	2.2	–	–	62
9–23 Ω	2M CrO_3 + 1M HF	rt	4.2	–	–	63
p, *n*	10 M HF + 0.5M HBr + 10 mM Br_2	rt	7.2	–	–	64
n, 10–100 Ω	3 M NH_4F + 1M $Cu(NO_3)_2$	22	31	2	20	65
Poly Si	0.25% HF + 35% HNO_3 + 15% H_2O_2	35	0.3	–	–	66
	8% HF + 75% HNO_3 + 17% CH_3COOH	rt	∼830	∼830	830	67
	1.3% HF + 49% HNO_3 + 49.7% CH_3COOH	70	65	52	–	68
p, 3 Ω	5% HF + 24% HNO_3 + 50% CH_3COOH	rt	–	200	–	69
p, 3 Ω	3% HF + 15% HNO_3 + 70% CH_3COOH	rt	18	9.7	–	70
n, $>8 \times 10^{18}$ cm^{-3}	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	rt	480[b]	–	–	71
n, $<10^{18}$ cm^{-3}	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	rt	27[b]	–	–	71
p, $>7 \times 10^{18}$ cm^{-3}	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	rt	330[b]	–	–	71
p, $<10^{18}$ cm^{-3}	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	rt	33[b]	–	–	71
p, *n*	Secco etch	rt	250	–	–	72
0.6–15 Ω	Schimmel etch	rt	292	–	–	73
1–3 Ω	Sirtl etch	rt	385	–	–	74
1–3 Ω	Wright etch	rt	100	–	–	74
	Yang etch	rt	250	–	–	75
	MEMC etch	rt	833	–	–	76
n, *p* 0.1–16 Ω	Sopori etch	rt	3500	–	–	77

Tab. 4 (*continued*)

Si[a]	*Solution*[c]	*T*	*(100)*	*(111)*	*(110)*	*References*
	Inorganic alkaline solutions					
	20% KOH (~4.2 M)	30	6	0.06	9	45
	20% KOH	100	560	19	930	45
	34% KOH	70	100	1.5	216	78
n, 1 Ω	23.4% KOH + 13.3% IPA	80	130	8	55	79
Poly Si	23.4% KOH + 13.3% IPA	80	38	–	–	79
n, *p*	0.5 M KOH	23	0.5	–	–	49
n, *p*	0.5 M KOH, stirred	23	6	–	–	49
n, *p*	0.1 M KOH, stirred	23	3.5	–	–	49
n	2 M KOH	rt	3.5	0.6	–	80
p	2 M KOH	rt	3.5	2	–	80
n, *p*	2 M KOH	70	50	–	–	49
p, 0–30 Ω	35% KOH, stirred	80	390	–	740	81
	44% KOH	80	1750	5.8	3500	82
	2 M LiOH	25	2	–	7.5	49
	2 M NaOH	70		–	14	49
	25% RbOH	50	21	–	33	83
n, 13 Ω	1.9 M NH_3	70	22	–	–	84
p	10% NH_4OH	70	10.7	–	–	85
p, 5 Ω	3.7 wt% NH_4OH, stirred	75	67	2.8	22	86
n, 13Ω	3×10^{-3} M H_2O_2 + 1.9 M NH_3	70	73	–	–	84
p, 10	0.5 M NH_4OH + 0.09 M H_2O_2	80	0.28	–	–	87
p, 10	5 M NH_4OH + 5 M H_2O_2	80	0.12	–	–	87
n, 0.05 Ω	2.65 M NH_3 + ~0.01M H_2O_2	80	333	–	–	88
Deposited Si	NH_4OH + H_2O (1:5), pH = 12.4	70	0.23[b]	–	–	89
Deposited Si	SC1	70	0.14[b]	–	–	89
p, *n*, 0.2–20 Ω	1:1 water:hydrazine	110	360	57	216	90
	48% hydrazine + 20% IPA + 32% H_2O	Boiling	50	–	–	91
	Organic alkaline solutions					
0.001–100 Ω	EDP, 8 ml H_2O + 17 ml $NH_2(CH_2)_2NH_2$ + 3 g $C_6H_4(OH)_2$	110	139	8.3	83	92
	B etch	115	69	5.2	–	93
	F etch	115	225	–	–	93
	S etch	115	170	4.2	190	45
p, 10–75 Ω	Etch 1	117	100	53	111	94
	25% TMAH,	70	45.3	1.5	89.9	54

(*continued overleaf*)

Tab. 4 (*continued*)

Si[a]	*Solution*[c]	*T*	*(100)*	*(111)*	*(110)*	*References*
n, 40 Ω	5% TMAH	90	225	5.4	308	95
n, *p*, 15 Ω cm	29% TMAH + 8.5% IPA	80	61	3.9	–	96

[a] cm^{-3} or cm Ω.
[b] Orientation not specified.
[c] HF solutions are generally in volume concentration while OH solutions are in weight concentration as the solvent water is not specified for its concentration.
Etch 1:100 g gallic acid + 317 ml ethanolamine + 144 ml water + 0.3 ml of 10% FC129 (3 M) surfactant.
SC1: $NH_4OH:H_2O:H_2O_2$ mixtures.
IPA: iso-2,2-propyl alcohol.
EDP or EPW: $NH_2(CH_2)_2NH_2 + C_6H_4(OH)_2 + H_2O$.
B etch: 7.5 ml $NH_2(CH_2)_2NH_2$ + 1.2 g $C_6H_4(OH)_2$ + 2.4 ml H_2O [93].
F etch: 7.5 ml $NH_2(CH_2)_2NH_2$ + 2.4 g $C_6H_4(OH)_2$ + 2.4 ml H_2O [93].
S etch: 7.5 ml $NH_2(CH_2)_2NH_2$ + 1.2 g $C_6H_4(OH)_2$ + 1 ml H_2O + 0.045 g pyrazine [93].
Sirtl etch: 1 48% HF: 1 5 M CrO_3 [73].
Secco etch: 2(49% HF):1(0.15 M $K_2Cr_sO_7$) at room temperature with ultrasonic agitation [72].
Sopori etch: 36:20:2 [HF (49%):CH_3COOH (glacial):HNO_3(70%)].
Wright etch: 60 ml 49% HF, 30 ml HNO_3 (69%), 30 ml 5 M CrO_3, 2 g $Cu(NO_3)_2.3H_2O$, 60 ml acetic acid (glacial), and 60 ml H_2O [74].
Yang etch: 0.75 M CrO_3 + 25% HF.
Schimmel etch: 0.33 M Cr^{6+} + 33% HF without ultrasonic agitation.
TMAH: tetramethyl ammonium hydroxide.
MEMC etch: 35 : 25 : 18 21(49% HF:70% HNO_3:HAc:H_2O) + 1g$(NO_3)_2.3H_2O$ per 100 ml of mixed solution [76].

to material and solution conditions, the etch rate of silicon in a given system depends on many operational parameters such as size and geometry of the sample, volume of the solution, stirring condition, ambient (light and air) control, and etching time. Thus, for a given silicon material and solution composition, etch rate may significantly vary when it is measured under different operating conditions. This is often responsible for the sometimes large difference between etch rates that can be found in identical systems from different investigations.

Table 5 shows the activation energies determined for the various etching systems. In general, the apparent activation energy, as determined from the dependence of etch rate on temperature, is 3–6 kcal mol^{-1} or 0.13–0.26 eV for diffusion-limited reactions, whereas it is 10–20 kcal mol^{-1} or 0.44–0.87 eV for surface-controlled reactions [1]. Using these criteria to evaluate the values of activation energy in Table 5 it appears that the etching process in each of the three etching systems may be under either diffusion control or surface reaction control depending on the specific solution composition and the operation conditions.

The etch rate of silicon relative to other materials is also important in device fabrications. The data on the relative etch rate, which is specific to a specific set of conditions, is widespread in the literature. Some examples are: $Si_{1-x}Ge_x$ in

Tab. 5 Apparent activation energy in various etching solutions

Si	*Solution*[c]	E_a[eV]	*References*
	Acidic solutions		
p(111), 3 Ω	5% HF + 24% HNO_3 + 50% CH_3COOH	0.42	69
n(111), 0.002 Ω	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	0.23	71
n(111), 5 Ω	0.42% HF + 17.5% HNO_3 + 67% CH_3COOH	0.54	71
p, *n*	5% HF + 63% HNO_3	0.17	97
n(100)	7 M NH_4F + 1 M $Cu(NO_3)_2$	0.52	65
n(111), 30 Ω	69.5 ml HNO_3 + 8 ml HF + 22 ml H_2O + 80 ml glacial acetic acid	0.21	98
p(111), 10 Ω	[CrO_3]/[HF] of 0.03–0.72 with 50 g CrO_3 in 100 ml H_2O and 48% HF	0.34–0	99
	91.5% H_3PO_4	1.15	100
	Inorganic alkaline solutions		
n(100), 5–50 Ω	25–45% KOH	0.56	101
(100)	42% KOH	0.59	45
(110)	42% KOH	0.60	45
(211)	6 M KOH, 50–90 °C	0.641	102
n(100), 1 Ω	23.4% KOH + 13.3% IPA	0.77	79
n(110), 1 Ω	23.4% KOH + 13.3% IPA	0.76	79
n(111), 1 Ω	23.4% KOH + 13.3% IPA	0.59	79
(100)	25% RbOH, 50–80 °C	0.47	83
(100)	3.7 wt% NH_4OH	0.22	86
(110)	3.7 wt% NH_4OH	0.19	86
(111)	3.7 wt% NH_4OH	0.19	86
(100)	1:1 water:hydrazine	0.28	90
(110)	1:1 water:hydrazine	0.41	90
(111)	1:1 water:hydrazine	0.25	90
	Organic alkaline solutions		
n(111), 1 Ω	EDP, 8 ml H_2O + 17 ml $NH_2(CH_2)_2NH_2$ + 3 g $C_6H_4(OH)_2$	0.48	92
(100)	B etch	0.47	92
(100)	EDP	0.40	45
(110)	EDP	0.33	45
(111)	EDP	0.52	45
(100), 10^{15} cm^{-3}	25% TMAH	0.52	103
(100), 2×10^{19} cm^{-3}	25% TMAH	0.64	103
(110)	22% TMAH, 60–90	0.39	95
(100)	22% TMAH, 60–90	0.61	95
(111)	22% TMAH, 60–90	0.40	95

[a] See Table 1 for solution compositions.
Activation energy: 1 cal mol^{-1} = 4.19 J mol^{-1} = 0.044 eV.

NH_4OH [85, 104], boron monophosphide epitaxial layer in $HF–HNO_3$ [105], SiO_2 in HF solutions [106], SiO_2 in KOH [82], SiO_2 and nitride in TMAH [96, 107], SiO_2 in NH_4OH [89], different oxides and metallic films [88], various glasses and silicides and metals in ethanolamine etchants [108], silicon oxide and metals in EDP [92], Al and SiO_2 in TMAH [95, 109], nitride and oxide in H_3PO_4 [100], SiO_2 in SC1 solutions [57], and SiO_2 in $HF + HNO_3 + H_2O_2$ [55].

The etching of silicon oxides is particularly important in silicon etching. Deposited silicon oxide serves as a dielectric layer, passive layer, or masking material for device fabrication and as a native oxide it is an essential part of the surface condition. Figure 16 shows the etch rate of different types of silicon oxides as a function of HF concentration, on which several general remarks can be: (1) etch rates of all the oxides increase with HF concentration, (2) the etch rates differ by several orders of magnitudes for different types of oxides with quartz being the slowest and anodic oxide being the fastest, reflecting the large difference in the structure of these oxides, (3) silicon as a solid is extremely stable in HF solutions compared to its oxide, as its dissolution rate is several orders of magnitude smaller than even that of quartz [3].

11.6.2 In Fluoride-based Solutions

11.6.2.1 Absence of Oxidants

The etch rate of silicon in HF solutions at room temperature at OCP is very low, in the order of 10^{-3} A s^{-1} in concentrated solutions (>25% HF) [52]. Figure 17 shows that dilution of concentrated HF results in an increase of etch rate [51, 110]. It also shows that the etch rate increases sharply from pH about 4, peaks at 6.2, and then decreases drastically to near zero at about 8.5 [53]. The etch peak has been attributed to the involvement of HF_2^- as the etching species, which shows a concentration peak at the pH about 6 [110]. In $HF–NH_4F$ solutions, the etch rate decreases with increasing HF at a constant NH_4F concentration and it increases with increasing NH_4F at constant HF concentration [111].

Etch rate depends on the time of immersion. For example, the etch rate of *p*-Si in concentrated HF decreases with immersion time from about 0.5 A s^{-1} at 1 minute to about 2×10^{-4} A s^{-1} at 10 days [52]. The relatively higher rate at the beginning is attributed to the etching of native oxide

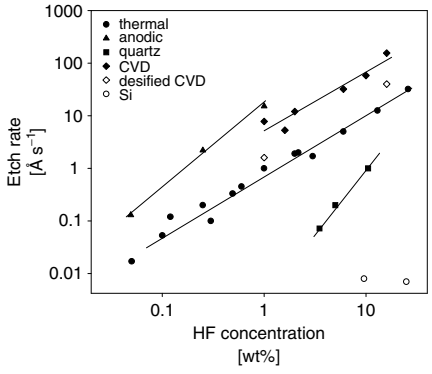


Fig. 16 Etch rate of different silicon oxides as a function of HF concentration (after Zhang [3]).

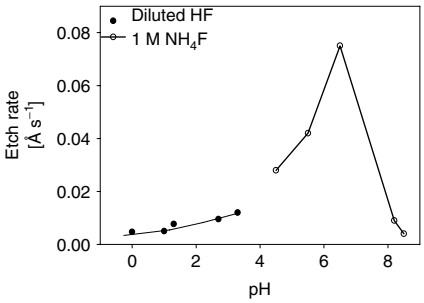


Fig. 17 Etch rate of Si in 48% HF at 25 °C as a function of pH by dilution (after Hu and Kerr [51]). Also, pH dependence of the etch rate of *n*-Si(III) in 1 M NH_4F (after Matsumura and Fukidome [110]).

and the lower etch rate at a longer time is due to the slow attack of OH^- ions on the relatively stable Si—H bonds.

The etch rate of silicon in HF does not have a clear dependence on dopant concentration. Also, oxygen in silicon has little effect on the etch rate in HF [51]. However, the reactivity of silicon in HF solution may be significantly increased when active elements such as sodium are present in the silicon [112]. Also, the etch rate of silicon has been found to increase with ion implantation, which produces an amorphous structure [113].

According to the reaction mechanisms discussed in previous sections, silicon dissolves chemically in KOH solutions by reacting to OH^- ions but electrochemically in HF solutions in which participation of charge carriers is required. In HF solutions at OCP, chemical dissolution rate is low because the concentration of OH^- is low. The electrochemical reaction is also low because there are few carriers at OCP. Thus, the increase of etch rate in HF with dilution can be attributed to the increase of hydroxyl concentration. The lack of direct proportionality between etch rate and OH^- concentration suggests that the rate depends on the surface concentration of OH^-. The relative contribution of chemical and electrochemical reactions depends on the pH [41].

Because of the involvement of charge carries, the dissolution rate in HF solutions increases with potential. The anodic dissolution consists of three regions with respect to potential. At potentials below the current peak, nonuniform anodic dissolution occurs, which is accompanied by significant hydrogen gassing, leading to the formation of a PS layer. At potentials above the current peak, uniform dissolution occurs with little hydrogen evolution, leading to electropolishing [3].

The etch rate in the electropolishing region is essentially the same for different doping conditions and orientations. The etch rate in the electropolishing region varies only with HF concentration and temperature. The presence of H_2SiF_6 has no effect on the etch rate in the electropolishing region [38]. Large etch rate can be obtained in HF-containing anhydrous organic solutions at high anodic potentials due to lack of oxide formation [114].

11.6.2.2 Effect of CrO_3 Addition

The HF—CrO_3 etching system is widely used for defect etching and delineation of junctions between silicon layers of different doping concentrations [75, 115]. The etch rate of silicon in pure HF solution is very low due to the lack of holes at OCP. Addition of CrO_3 increases the etch rate due to the increase of surface hole

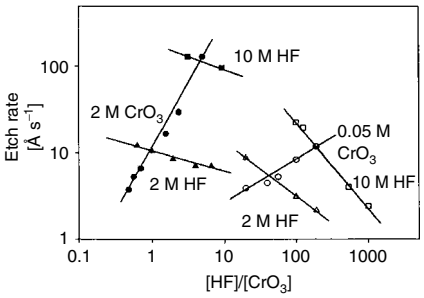


Fig. 18 Etch rate as a function of the ratio of HF and CrO_3 concentrations ([HF]/[CrO_3]) (after Gaspard et al. [62] and Meerakker and Vegchel [63]).

concentration resulting from the reduction of Cr^{6+}.

According to Meerakker and Vegchel [62, 63], the etching process has different kinetics at high and low HF/CrO_3 ratios as shown in Fig. 18. For etching in CrO_3–HF solutions with high HF/CrO_3 ratios (>20) the rate equation can be expressed as

$$r = 1.91[CrO_3](1 + 0.23[HF]^{1.5})/(4 + 0.23[HF]^{1.5}) \quad (1)$$

where r is in micrometers per minute and concentrations are in molar. At low HF/CrO_3 ratios (<10) the rate equation is

$$r = [CrO_3][HF]^{1.5}/(38[CrO_3] + 1.8[HF]^{1.5}) \quad (2)$$

On the other hand, Chu and Gavaler [116] found that the dissolution rate of silicon in $HF-CrO_3$ solutions is second order with respect to the concentration of HF and first order with respect to $Cr_2O_3^{2-}$ ions. Heimann [99] suggested that the etch in CrO_3-HF-H_2O solutions follows a two-step reaction mechanism. In the first step, silicon is oxidized to form a silicon suboxide SiO_x ($0.67 < x < 1$). The second step consists of the dissolution of the silicon oxide by HF.

11.6.2.3 Effect of HNO_3 Addition

The $HF-HNO_3$ etching system is perhaps the most used isotropic etchant for silicon [117, 118]. The mixture of these two concentrated acids can be diluted with water or any other dilutant to give a wide range of etch rates as represented by the triaxial plot shown in Fig. 19 [119]. The etch rate with respect to composition of the etchant shows two extreme modes of behavior. In the region of high nitric acid composition etch rates are only a function of HF concentration while in the region of high HF concentrations the nitric acid concentration determines the etch rate. The etch rates decrease with addition of a dilutant but are little affected in the two extreme regions. As a dilutant, acetic acid acts qualitatively the same as water [119]. However, acetic acid is preferred because it results in less dissociation of the nitric acid, which preserves the oxidizing power of the HNO_3 for a wider range of dilution than water [117]. The etching in $HF-HNO_3$ solutions is largely isotropical with etch rate of (100) surface very close to that of (111) surface as shown in Table 4.

The $HF-HNO_3$ etching system was systematically investigated by Schwartz and Robins [97, 119]. They found that in the low HF concentration region, the etching process is a diffusion-controlled process; etch rates of different silicon materials in

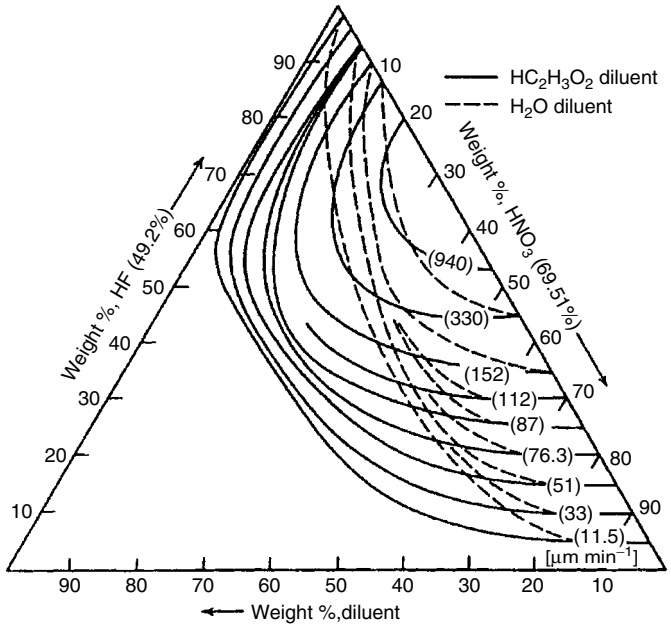


Fig. 19 Iso-etch curves for silicon in HF:HNO_3: diluent system (after Robbins and Schwartz [119]).

this concentration region are essentially the same. In the high HF region, two different values of activation energy are observed, one in the range of 10–14 kcal mol^{-1} at higher temperatures and the other in the range of 20 kcal mol^{-1} at lower temperatures; both are characteristic of a surface-controlled process. Consequently, dopant concentration, crystal orientation, and defects play an important role in the etching process. Addition of a dilutant results in an activation energy that is between these two ranges.

The etching process of silicon in HF–HNO_3 solutions is an electrochemical process in which the dissolution of silicon takes place at local anodic sites while the oxidizing agent is reduced at local cathodic sites [120]. The etching process may also depend on surface carrier concentration determined by doping since lowly doped materials etch slower than highly doped materials [65]. The reduction of HNO_3 has an effect that changes the surface concentration of the carriers. The etching, with a rate about the same for both *n*-Si and *p*-Si, proceeds mainly through capturing the holes, which are injected into the valence band due to the reduction of HNO_3 [118, 121].

According to Schwartz and Robins [97, 119], the etching of silicon in the HF–HNO_3 system follows a chemical process with two basic reaction steps. In the first step, silicon is oxidized by HNO_3, which is followed by the dissolution of the oxidized Si by HF. The overall reaction is

$$3Si + 12HF + 4HNO_3 = 3SiF_4 + 4NO + 8H_2O \quad (3)$$

This reaction involves several intermediates. Because addition of trace amounts of $NaNO_2$ increases the etch rate in the solutions it is considered that HNO_2 is produced and acts as an active oxidizing agent. A steady-state condition can be achieved only after a concentration of HNO_2 in excess of a threshold value is built up in the vicinity of the surface either through the reduction of HNO_2 at a sufficiently high rate or by addition of small amounts of $NaNO_2$ in the solution.

11.6.2.4 Addition of Other Oxidants

The etching process in Br_2 containing HF solutions, unlike that in the HNO_3 solutions, is mainly of a chemical nature [64]. At OCP, Si reacts directly with Br_2 without involving charge carriers in the bands. The reduction of Br_2 at cathodic potentials is facilitated by the electrons from the conduction band. The etch rate, which is independent of doping type and surface orientation, is limited by the mass transport of the Br_2 in the solution. Thus, for *p*-Si, the etch rate is almost constant at cathodic potentials at which no electron is present on the surface. On the other hand, the etch rate of *n*-Si decreases to zero at cathodic potentials because Br_2 is efficiently reduced with the conduction band electrons.

The reduction of H_2O_2 on silicon in HF solutions is similar to Br_2, being primarily a conduction band process with very little hole injection into the valence band [122]. The etch rate is 0.1–0.3 A s^{-1} in 10^{-3} to 1.3 wt% NH_3F with 5% H_2O_2 [57]. It increases rapidly with increasing fluoride concentration.

11.6.3 Alkaline Solutions

The most commonly used alkaline solutions have been KOH and EDP (or EPW) which is a mixture of ethylenediamine (ED or E), pyrocatechol (P), and water (W) [45, 84]. Other solutions such as $NHOH_4$, hydrazine, ethanolamine, and tetramethyl ammonium hydroxide (TMAH) are also used. The development and application of these etching solutions are dictated by a number of factors such as etch rate, anisotropic selectivity, corrosiveness to masking materials, surface quality, processing controllability, safety, and, more recently, environmental impact.

The advantages of KOH solutions are its high (100)/(111) etch rate ratio and nontoxic nature [82]. The disadvantages are the significant etch rate of SiO_2, undesirable etching of metallic film, and potential potassium contamination [103, 123]. KOH solutions also lack the controlled uniformity and boron specificity as do the EDP solutions [124]. On the other hand, EDP has several advantages: (1) it is anisotropic, (2) it can be masked by a variety of materials such as SiO_2, Si_3N_4, Cr, and Au, (3) it exhibits near-zero etch rates on heavily boron doped silicon [123]. However, EDP solutions have much lower anisotropic etch rate ratios than KOH. Also, like KOH, EDP etches aluminum rapidly and EDP solutions have to be controlled at a specific temperature range to avoid precipitation. In addition, EDP is carcinogenic [124]. Because of the disadvantages of these two etching solutions, other systems have been developed [83, 90, 108, 125].

Silicon etching in KOH solutions have been extensively investigated, resulting in a body of information that shapes the current understanding of the etching behavior of silicon in alkaline solutions. The major characteristics and the principal reaction processes involved in all alkaline solutions appear to be similar to that in the KOH system although the detailed characteristics vary from system to system. Most notably,

all alkaline solutions show the sensitivity of etch rate to crystal orientation, which is the basis for anisotropic etching. Also, all these etchants show an etch rate reduction for highly boron doped materials. The alkaline etchants can be made of organic and inorganic solutions, but all of them appear to require the presence of water to etch silicon at significant rates.

11.6.3.1 KOH Solutions

The etch rate of silicon in KOH may vary from as low as 1 A s^{-1} in dilute KOH (e.g. 0.5 M) at room temperature to as high as 2000 A s^{-1} in a concentrated solution (e.g. 40% KOH) at high temperatures. Figure 20 shows that the etch rate of (100) silicon in KOH solution at room temperature increases with concentration, peaks at about 5 M, and then decreases with further increasing concentration [49]. Agitation of the solution can significantly increase the etch rate, indicating that mass transport is involved in the rate-determining processes. A similar etch rate dependence on concentration is found at a temperature of 80 °C but unlike at lower temperatures solution stirring has little effect on the etch rate [45, 79]. KOH solutions are often buffered with isopropyl alcohol (IPA) [79, 126]. Addition of IPA into KOH solution results in a decrease of etch rates. Also, etch rate tends to decrease with etching time; a decrease of a factor of 3 is observed over a 24-hour period [127].

The most characteristic feature of etching in KOH solution is its anisotropic nature, that is, etch rate varies with the crystal orientation of the silicon wafer. The etch rates for the major crystal planes are {110} > {100} > {111} [82]. The exact difference between the etch rates of these planes depends on concentration, temperature, and measurement procedure.

For a given type and orientation, the etch rate is essentially independent of the doping concentration up to a concentration of about 10^{19} cm^{-3}, above which the etch rate drastically decreases with further increasing doping concentration, particularly for boron doped materials [33, 80]. Doping type has little effect on the etch rate of (100) surface but a significant effect on that of (111) surface [18].

The etch rate of all silicon materials in KOH depends, to a varying extent, on the potential [18, 128, 129]. Figure 21 shows the effect of potential on the etching rate in a 2 M KOH solution at room temperature. The etch rate is the highest at OCP. At potentials positive of the passivation, potential etching stops because of the formation of oxide film. At cathodic potentials, the effect is different

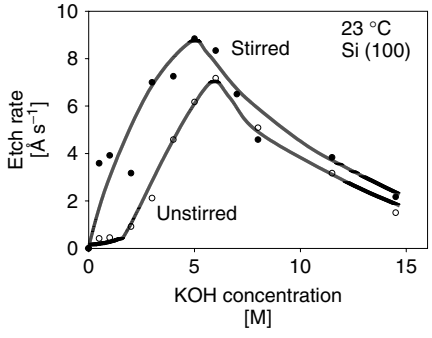


Fig. 20 Etch rate versus molarity for 23 °C on the (100) surface. Data for *n*- and *p*-type Si samples were virtually the same (after Glembocki et al. [49]).

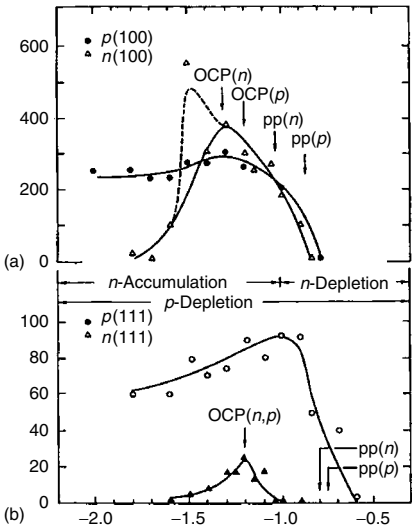


Fig. 21 The potential dependence of the etch rates for (a) $p(100)$ and $n(100)$, and (b) $p(111)$ and $n(111)$ low-doped Si. The data have an experimental spread of approximately ±20%. The open-circuit (OCP) and passivation (pp) potentials are marked. Note that the OCP is the same for the $n(111)$ and $p(111)$ samples (after Glembocki et al. [18]).

for n-Si and p-Si. For n-Si, the etch rate decreases with increasing cathodic bias and etching stops at certain potentials. For p-Si, the etching maintains relatively high rates comparable to that at OCP.

The difference between n-Si and p-Si at the cathodic potentials indicates the effect of carriers on the etching process. At cathodic potentials, hydrogen evolution may obtain the electron either directly from the dissolving surface silicon atom or from the semiconductor. For n-Si, cathodic bias provides a high concentration of surface electrons for hydrogen reaction resulting in a decrease of the dissolution rate. However, for p-Si, electrons are the minority carriers and thus the etch rate remains more or less constant at cathodic potentials.

The morphology of the etched surface depends on the composition of the solution. In general, because of the anisotropic nature of the etching of the crystal confined in a given area, perfectly smooth lateral (111) sidewalls form on both (100) and (110) wafers [130, 131]. At low KOH concentrations, below 30%, there is a tendency to form pyramidal islands on (100) surface. The density of hillocks varies with temperature, potential, and the presence of chemical agents in the solution [34, 47].

11.6.3.2 **Reaction Kinetics**

The etching of silicon in KOH is a process of dissolution and hydrogen evolution. Two hydrogen molecules are evolved for each silicon atom dissolved during the etching [47]. The main etching product is determined to be $SiO_2(OH)_2^{2-}$ which can further be polymerized [48, 132]. OH^- is an active etching species since OH^- concentration decreases during etching. Water is also an active species because silicon does not etch in water-free alcohol that contains KOH [47]. K^+ as a cation present in the solution is not directly involved in the etching reaction because the etching behavior of silicon is essentially similar in different alkali solutions [33, 49]. In IPA buffered solutions, the IPA does not alter the composition of the solutions or participate in the etching reactions [48, 132]. The activation energy of the etching process is 0.56–62 eV for KOH of different concentrations indicating that the etching is controlled by a surface process [45].

The dependence of etch rate on KOH concentration in the concentration range of 10 to 60% has been found to best fit with

an equation [45]

$$R = k[H_2O]^4[OH^-]^{\frac{1}{4}} \qquad (4)$$

A similar rate equation has also been proposed considering only the activities of H_2O and OH^- [49]:

$$R = C[H_2O_{free}]^r[OH^-_{free}]^s \qquad (5)$$

According to this model, at low hydroxide concentrations, $[OH^-_{free}]$ is low resulting in a low etch rate, while at very high concentrations $[H_2O_{free}]$ is also low resulting in low etch rates. In moderately concentrated solutions, both $[H_2O_{free}]$ and $[OH^-_{free}]$ are high and etch rate exhibits a maximum. The model is, however, not applicable for low concentrations from 0.1 to 2 M due perhaps to a rate-limiting reaction associated only with free water [49].

11.6.3.3 Other Inorganic Solutions

The etching processes in other alkali hydroxides are found to be, in general, similar to that in a KOH solution, namely, free H_2O and OH^- are the active agents. The difference among the solutions such as NaOH, LiOH, NH_4OH, RbOH, and CsOH is mainly due to the difference in the hydration number of these solutions [36, 49, 83, 133].

Among the inorganic alkaline solutions, NH_4OH is special in two aspects. First, when mixed with H_2O_2, it is widely used as a cleaning solution (known as *SC1 solution*) for cleaning organic and metallic contaminants on the silicon surface. Secondly, it has the characteristics of both inorganic alkaline etchants, such as KOH, and organic etchants, such as EDP. It has a low etch rate for thermal SiO_2 giving an etch rate ratio of Si/SiO_2 of 1:8000 [88]. Also, as a chemical NH_4OH has much lower heavy metal concentrations compared to other hydroxides [134]. However, NH_4OH tends to decompose at elevated temperatures, especially with more concentrated solutions [84].

The etch rate in NH_4OH solutions increases with concentration from about 45 A s^{-1} at 1 wt% to a peak of about 80 A s^{-1} at 10 wt% [86]. For a given concentration, the etch rate increases with pH [84]. The etch rate is dependent on potential with the peak etch rate of both *p*-Si and *n*-Si occurring near OCP [135]. The etch rate of *p*-Si decreases but maintains a definite value at cathodic potentials while that of *n*-Si decreases to zero at large cathodic potentials.

Figure 22 shows the etch rate in NH_4OH as a function of H_2O_2 concentration [84]. A sharp drop of etch rate to nearly zero

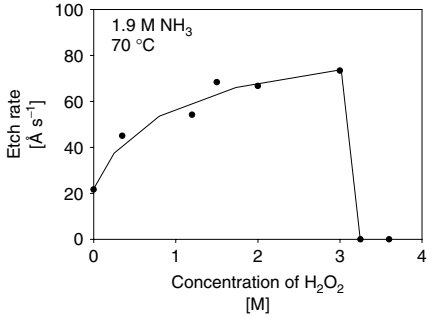


Fig. 22 Etch rate as a function of the H_2O_2 concentration in a solution of NH_3 (after van den Meerakker and van der Straaten [84]).

is observed for concentrations higher than 3.2×10^{-3} M, which is due to the passivation by formation of silicon oxide on the surface. In solutions containing more than 3.2×10^{-3} M H_2O_2, the H_2O_2 as an oxidant polarizes the silicon surface to a potential value ($-0.4\ V_{sce}$) above the passivation potential. The etching of silicon in solutions with a H_2O_2 concentration below the critical value has a chemical component and an electrochemical component due to the reduction of H_2O_2. The relative contribution of the two components changes with the concentration of H_2O_2 [136].

In $NH_4OH-H_2O_2$ solutions at a given H_2O_2 concentration, the etch rate increases with increasing NH_4OH concentration and reaches saturation value at a certain concentration [87]. The saturation etch rate, reflecting the dissolution rate of the passive oxide film, decreases with increasing H_2O_2 concentration, reaching a limiting value of about 0.12 A s^{-1} at high concentrations.

Formation of pyramidal hillocks bounded by (111) crystal planes on (100) wafer is a problem in NH_4OH solutions for all concentrations, temperatures, stirring conditions, and sample preparations. The formation of hillocks significantly reduces the etch rate in the (100) direction [88]. Addition of H_2O_2 in NH_4OH solution has an effect of suppressing the formation of hillocks.

11.6.3.4 Hydrazine

Hydrazine, N_2H_4, may be thought of as derived from NH_3 by replacing an H atom by the NH_2 group. A hydrazine–water (1:1) solution shows some desirable characteristics [90]. The etch rate of oxide, nitride, and most metals in hydrazine solutions is very low. No precipitation occurs during etching and the etching produces good surface quality and sharply defined corners. This etchant, however, has a low anisotropic selectivity and is also toxic.

A 1:1 hydrazine solution at 110 °C gives an etch rate of 360 A s^{-1} and a (100)/(111) ratio of about 6 [90]. Anodic polarization of silicon electrode in hydrazine shows i–V curves to be similar to those in other alkaline solutions [137]. The etch rate is the highest when the mixture contains no IPA indicating that IPA acts as a dilutant but not as a complexing agent. Some mixtures may result in stains on the surface, while other solutions may result in the formation of an oxide film, which stops the etching. The slowest etching planes are the (111) planes and the fastest planes are (211) planes. The etched surfaces may vary from being highly faceted to very flat depending on water content and temperature. The faceted surfaces are made up of square pyramidal hillocks, and the flat surfaces are marked by a cellular structure.

11.6.4 Organic Solutions

11.6.4.1 EDP Solutions

EDP or EPW is a mixture of ethylenediamine (ED or E), pyrocatechol (P), and water (W). This solution usually operates at 110–120 °C (about boiling point) [93]. The typical compositions and etch rates can be seen in Table 5. Figure 23 shows the etch rate as a function of the water content [92]. No etching occurs in solutions without the presence of water and the maximum etch rate is observed with a water molar fraction of about 0.6. The etch rate increases with increasing pyrocatechol content to about 5 mol% above which etch rate becomes constant. There is a definite etch rate without addition of pyrocatechol,

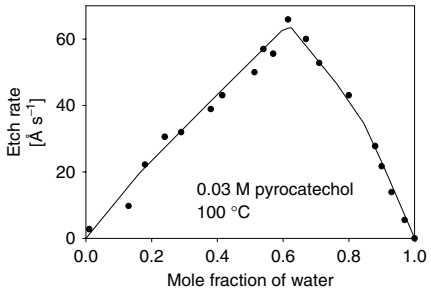


Fig. 23 Etch rate of silicon as a function of the water content of the etch with pyrocatechol content held constant (0.03 M) (after Finne and Klein [92]).

suggesting that it is not an essential component of the etching system but rather a catalyst. Addition of other agents such as pyrazine and *p*-benzoquinone has a similar effect [93].

EDP etch solutions can be affected by many operation factors such as bath design, contamination level, silicon content in the solution, and oxygen pressure [93]. Diffusion plays an important role in the etching of silicon in EDP [45]. Trace amount of chemical agents such as diazine, pyrazine, and quinone changes the etch rate and the difference in the etch rates among the crystal planes [46]. The etch rate changes significantly when exposed to air due to the effect of dissolved oxygen. Also, insoluble residues and rough raindrop-like patterns tend to form at low temperatures and with aged solutions [93, 112]. Residue formation occurs with certain compositions even at high temperatures.

The effect of potential is shown in Fig. 24; etch stops at passivation potential and decreases with cathodic polarization for both *n*-Si and *p*-Si [138, 139]. The etch rate dependence on potential is rather different from that in KOH in which *p*-Si etch rate varies only slightly with cathodic polarization (Fig. 21). Also, etch rates at OCP as shown in Fig. 24 are much smaller than the peak values indicating that electrochemical processes play a more important role than that in KOH. Etch rate in EDP is independent of illumination [46].

11.6.4.2 **Etching Reactions**

The presence of water is essential for anodic current and etching to occur in EDP [93, 139]. Water is the active etching species since no etching occurs

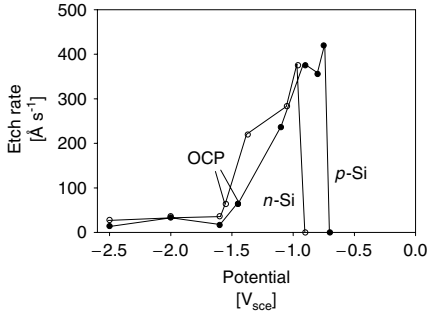


Fig. 24 The effect of potential on etch rate of silicon in EDP etchant (after Gealer et al. [138]).

for either anhydrous amine or anhydrous amine-pyrocatechol mixtures [92]. Etching does not occur without water even in the presence of fluoride ions. The reaction is mainly a chemical mode with about 20% electrochemical contribution, which is considerably higher than that in KOH solutions [19]. The ratio of the evolved hydrogen to dissolved silicon is close to 2, meaning that evolution of hydrogen is the only reduction reaction accompanying the oxidation of silicon. The dissolution product is $SiC_{22}N_4H_{30}O_6$ [93].

The etching activation energy is about 0.48 eV in a solution with a ratio of E:P:W of 28:3:69% suggesting that the etching process is probably not controlled by diffusion [92]. Etching occurs by oxidation of the silicon by water and hydroxyl ions resulting in evolution of hydrogen and hydrous silica at the silicon surface. The hydrous silica is removed through the formation of the amine-soluble silicon(IV). The overall reaction process was suggested by Finne and Klein [92] as follows

$$NH_2(CH_2)_2NH_2 + H_2O \longrightarrow NH_2(CH_2)_2NH_3^+ + OH^- \text{ Ionization} \quad (6)$$

$$Si + 2OH^- + 4H_2O \longrightarrow Si(OH)_6^= + 2H_2 \text{ Oxidation} \quad (7)$$

$$Si(OH)_6^= + 3C_6H_4(OH)_2 \longrightarrow [Si(C_6H_4)_3]^= + 6H_2O \text{ Chelation} \quad (8)$$

According to Finne and Klein, an amine environment is required for the pyrocatechol to dissolve (chelate) the hydrous silica. Different amines have different effects on the etching due to their protophilic activity. The alkalinity of EDP medium is weak in comparison to KOH solutions. The low alkalinity of EDP medium does not allow the formation of a soluble silicate and the solvation of silicon is provided by chelating with pyrocatecholate that exists as a partially dissociated monoanion but further dissociates on complexation with Si(IV) [139].

11.6.4.3 Ethanolamine

Ethanolamine ($NH_2CH_2CH_2OH$) based anisotropic etchants are developed aiming to replace the toxic EDP etchant [124, 140]. Linde and Austin [124] reported a functional formula which consists of 100 g gallic acid + 305 ml ethanolamine, $NH_2CH_2CH_2OH$, +140 water + 1.3 g pyrazine + 0.26 ml FC129 (3 M) surfactant. The typical etch rate of (100) silicon in this etchant at 118 °C is about 300 A s^{-1}. It has high boron selectivity and good uniformity. Also, the etch rate ratios of silicon to other materials such as carbide, nitride, oxides, and most metals in ethanolamine etchants are very large [108].

Gallic acid functions as an effective silicon complexing agent in aqueous ethanolamine solutions. In the absence of it the etch rate is zero. Etch rate increases with water content. The additions of pyrazines, pyridazines, and triazoles show various catalytic effects on the etching process [140]. The catalysts that lead to faster oxidation result in faster etch rates, and the difference among the catalysts is attributed to a steric effect. Oxidative catalysts tend to influence the etching selectivity of the major crystal orientations [94].

11.6.4.4 Tetramethyl Ammonium Hydroxide (TMAH)

TMAH, $(CH_3)_4NOH$, etchant has been developed more recently for anisotropic etching of silicon [59, 109, 141]. It is relatively safe to use and presents no special

disposal issues. It has very low oxide etch rate and does not attack aluminum if the solution contains a certain amount of silicate. However, hillocks tend to develop in this solution [95, 109].

The etch rate increases with TMAH concentration up to about 4% above which it becomes relatively constant [137]. The etch rates of (100) and (110) planes decrease as a function of TMAH concentration at higher concentrations [95]. The etch rate tends to increase with increasing amount of dissolved silicon in the solution [109]. The etch rate ratio of (100)/(111) planes varies with TMAH concentration and temperature, with a minimum at about 25% TMAH [95]. The slowest etching plane is (111), 1.5 A s^{-1}, and the fastest one is (210), 99.5 A s^{-1} [54]. Addition of IPA into the TMAH solution slows down the etching [96]. Addition of pyrazine into the TMAH solution increases the etch rate slightly and improves surface smoothness [103].

Illumination during etching has no effect on the etching rate [59]. The etch rate decreases sharply at anodic potentials corresponding to the occurrence of passivation [142]. The etch rates at OCP are close to the peak values which is similar to the situation in KOH, indicating very little contribution from the electrochemical process at OCP. At cathodic potentials, *n*-Si and *p*-Si are different; while the etch rate on *p*-Si remains relatively unchanged, that of *n*-Si decreases with increasing cathodic polarization, which is also similar to that in KOH. The dissolution products in 2 and 20% TMAH solutions are mainly of $SiO_2(OH)_2^{2-}$ and its polymers [143]. The etching mechanism has been suggested to be similar to that in KOH.

11.7 Etch Rate Reduction of Heavily Doped Materials

For a given type and orientation, the etch rate of silicon in alkaline solutions is largely independent of doping concentration up to a concentration of about 10^{19} cm^{-3} [33, 80]. At a doping level of about 2×10^{19} cm^{-3}, the etch rate of boron doped silicon drastically decreases with increasing dopant concentration [33, 86, 144]. Reduction by as much as 3 orders of magnitude can be obtained by varying the boron concentration from about 10^{19} to above 10^{20} cm^{-3}. This feature has been widely used as an etch-stop technique for the fabrication of silicon microstructures.

Doping with other elements such as carbon, germanium, and phosphorus also results in etch rate reduction as shown in Fig. 25 [33, 145]. Ion implantation of

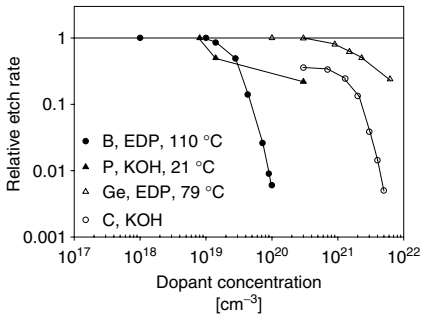


Fig. 25 Normalized <100> silicon etch rates as a function of dopant concentration for boron, phosphorus, germanium, and carbon in KOH and EDP solutions (after Seidel et al. [33] and Lehmann et al. [79]).

species such as arsenic to the level of 10^{17} cm^{-3} has also been shown to reduce the etch rate [146]. On the other hand, doping with As, P, and Sb was found to change little from 10^{14} to 10^{20} cm^{-3} in 23.4% KOH + 13.3% IPA [79].

The occurrence of etch rate reduction on highly boron doped materials appears to be independent of doping methods, whether by solid-source diffusion, epitaxial growth, or ion implantation [147]. However, the boron concentration at which significant reduction occurs is different for different methods of doping. The critical boron concentration for etch rate reduction to occur is affected by the defect density in different doped materials. For similar boron concentrations, the amount of etch rate reduction in KOH solutions decreases with increasing defect density.

Etch rate reduction is observed on highly boron doped silicon in all major alkaline etching solutions [33, 80, 103]. In NH_4OH etch rate reduction is stronger than in EDP and KOH. Much larger etch rate reduction, as high as 8000 times, has been observed [86]. Etch rate reduction is not observed for highly doped silicon in HF-based solutions [3].

11.7.1 Mechanism

Several models have attributed the phenomenon to increased lattice distortion and defect density, surface passivation, or shortened electron life. In an earlier investigation, it was suggested that etch rate reduction is caused by the increased lattice distortion from high doping. Later studies indicated that the reduced etch rate of high boron doped material is not likely to result from lattice distortion or stress.

The shortened electron life model is based mainly on the observation that the decrease of etch rate with boron concentration exhibits an inverse fourth power dependence on the boron concentration [33, 46, 147]. The etch rate reduction at high boron doping levels is attributed by Raley et al. [46] to the reduced electron concentration required for the reduction of hydrogen in the etching process. The rate of reduction reaction is a fourth order reaction with respect to electron concentration n

$$R \propto n^4 \tag{9}$$

Seidel et al. [148] proposed a similar model, ascribing the reduced etch rate of highly doped material to the reduced electron concentration caused by the degeneracy induced by the heavy doping [33]. According to this model, the etch rate decreases at a concentration corresponding to a doping level of 2.2×10^{19} at which the material is degenerated and the Fermi level moves into the valence band, making silicon behave like a metal. The fourth power decrease of etch rate is explained by the fact that for the dissolution of one silicon atom, four charges are required to transfer across the solid/electrolyte interface.

The shortened electron life model could not explain the etch stop on n-Si, in which there are abundant electrons in the conduction band. Also, this model assumes electrochemical reactions responsible for the etching process in alkaline solutions, in which they are in fact largely in chemical nature.

On the other hand, the passivation model, proposed by Palik et al. [80, 149], attributed the etch rate reduction to the easier formation of an oxide film on highly doped silicon. It is supported by a number of experimental observations. First, the difference between passivation potential and OCP decreases with increasing doping concentration, implying easier passivation

for highly doped materials. Secondly, elipsometric measurements of the samples during etching in KOH solutions indicate the existence of a surface phase on highly doped materials but not on lightly doped materials. This coincides with the i–V curves of highly doped materials, which show no current peak, implying that an oxide film may already exist near OCP. Because of the relatively slow etch rate of oxides in alkaline solutions, boron oxide, and hydroxides, once formed tend to stay on the surface and would thus passivate the surface and inhibit further dissolution of the silicon.

Palik et al. [80, 146] further suggested that the tendency to grow an oxide layer on highly doped materials is due to the strain in the silicon surface, which enhances the Si—B bond such that a borate glass is formed. The strong B—Si bond tends to bind the lattice more rigidly, thereby increasing the energy required to remove a silicon atom high enough to stop etching. The atomic concentration of boron doped silicon at etching stop levels corresponds to an average separation between boron atoms of 20–25 Å and is near the solid solubility limit (5×10^{19}) for boron substitutionally introduced into the silicon lattice. Silicon doped with boron is under tension as the smaller boron atom enters the lattice substitutionally, thereby creating a local tensile stress field.

11.8 Anisotropic Etching

Anisotropic etching, that is, different dissolution rates on different crystal planes, is a characteristic feature of silicon etching in alkaline solutions. Strictly speaking, the etch rate of silicon always depends, to a various extent, on crystal orientation in all etching solutions, acidic or alkaline. However, the etch rate difference among different planes is small in acidic HF solutions compared to those in alkaline solutions. Figure 26 shows the etch rate ratios of (100)/(111) and (110)/(111) planes in various solutions.

In HF-based solutions the ratios are close to 1, varying only slightly with different solution compositions. An exception is in NH_4F solution with the addition of copper ions in which the etch rate ratio of (100)/(111) is found to be as high as 15 [65]. Addition of silver or gold ions does not have such an effect. The special effect of copper ions on anisotropicity may indicate that the copper deposition

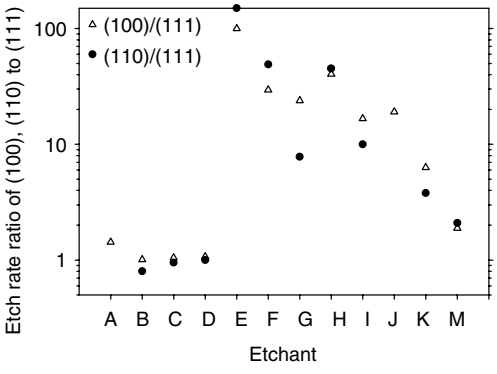


Fig. 26 Etch rate ratio of (100)/(111) and (110)/(111) in different etching solutions (data from Table 6).

process may be anisotropic since etching in solution with metal ion oxidants is accompanied by deposition of the metal on the surface.

The etch rate ratio in alkaline solutions, on the other hand, tends to be large and to greatly change with solution composition and etching conditions. KOH-based etching solutions are the most widely used anisotropic etchants [2, 45, 82]. These etchants show the highest etch rate ratio of (100)/(111) and (110)/(111) planes. Ratios as high as 300–400 for (100)/(111) and 600 for (110)/(111) can be obtained [150, 151]. The etch rate ratio is even higher at room temperature; ratio as high as 2500 is achievable for (110)/(111) planes in KOH solution [151]. Addition of IPA into KOH results in a decrease of etch rates differently for the three major orientations [79]. As a result, at certain alcohol concentrations, the order of the etch rates is changed due to the reverse order of the etch rates for (100) and (110).

Among the organic alkaline systems, solutions consisting of ethylenediamine, water, and pyrocatechol (EDP) are among the most widely used. The etch rate ratio in EDP, although lower than that of KOH, is rather high compared to other organic etchants as shown in Fig. 26. Etch rate ratios are almost constant with respect to doping concentration up to the heavily doped levels where etch rate reduction occurs [45, 80, 92]. A highly defective surface, such as a saw-damaged surface, may not exhibit anisotropic etching rates [152].

The (111) surface is the slowest etching plane in alkaline etchants in all reported investigations [45, 54, 91]. The fastest etch plane in alkaline solutions, on the other hand, depends on solution composition. For example, for silicon surfaces of different orientations ranging from 0 to 45° from the (010), the plane of fastest etch rate is (210) in 35% KOH at 80 °C [81]. In the etching system of hydrazine and water, and hydrazine, IPA, and water, the fastest plane is (211) [91, 131]. In KOH, (411) is found to be the fast etching plane in one study [153] while (320) in another [154], (331) in KOH–propanol–water system [52], and (110) in CsOH [154]. Table 6 shows that (210) and (310) are the relatively fast etching planes in TMAH solution while (110) and (320) are the ones in the

Tab. 6 Etch rates of different crystallographic planes in 25% TMAH and 34% KOH at 70 °C, in A s^{-1} [54, 78]

Orientation	*TMAH*		*KOH*	
	Etch rate	*Ratio of X/(111)*	*Etch rate*	*Ratio of X/(111)*
(100)	45.3	30.2	100	67
(110)	88.7	59.1	216	144
(210)	99.5	66.3	212	141
(211)	87.8	58.5	151	101
(221)	79.5	53	94	63
(310)	98.2	65.5	181	121
(311)	96.0	64	179	119
(320)	95.7	63.8	220	147
(111)	1.5	1.0	1.5	1.0

KOH solution [54, 78]. For (110) and (100) planes, which one etches faster depends on the formulation of the etchant as shown in Fig. 26. It is interesting to note that in HF solutions the etch rate of (111) plane may not be the lowest as shown in Table 4.

The etch rate in the vicinity of {111} crystal planes is extremely sensitive to small angular misalignments as shown in Fig. 27 [45, 54, 92]. Deviation by 1 °C from the (111) plane leads to an increase of etch rate by a factor of 5 in EDP and TMAH and by about a factor of 2 in KOH. Because of the sensitivity of the etch rate to small deviation from the (111) planes, etch rate ratios may be much reduced for slightly misorientated (111) samples relative to that of perfectly oriented (111) material. The etch rate ratio of different planes is very sensitive to the misalignment of mask, which is considered to be responsible for the large variations sometimes reported in the literature on etch rate ratios under identical conditions [82, 155]. It has been shown that for a misorientation of half a degree in the alignment of the mask, the etch rate ratio of {110}:{111} in 44% KOH at 85 °C changes from above 400 to 70.

11.8.1 Rate-Limiting Process

Etching anisotropicity arising from differences in the atomic arrangement on the surface of different orientations necessarily indicates that the process is controlled by lattice-dependent surface reactions. Any shift from surface-controlled processes to nonsurface processes will result in a decrease of anisotropicity. Experimental data generally show that although mass transport may be rate limiting for the fast etching of (100) plane in some solutions, it is the surface-controlled processes that are the rate-limiting steps in the slow etching of (111) plane in all anisotropic etchants.

The rate-controlling process can also be evaluated by the apparent activation energy as shown in Table 5. Etch rate ratio appears to increase with increasing activation energy, with the highest ratio and activation energy being in KOH. High activation energies are also found in $HF-HNO_3$ solutions with high HF concentrations, which are isotropic etchants. In this case, the rate controlling surface process is associated with the surface carrier concentration instead of the lattice structure as in alkaline solutions. An important difference between isotropic etching and anisotropic

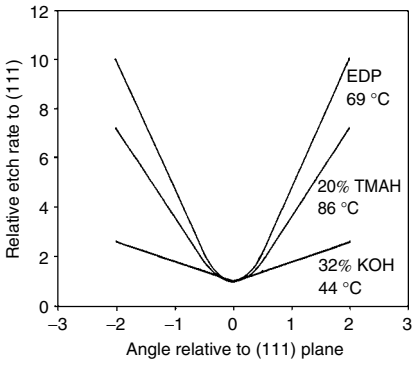


Fig. 27 Etch rate as a function of the angle of deviation from the (111) plane in three different alkaline solutions (data from Refs 45, 54, 92).

etching, when both are controlled by surface processes, is that the former is not sensitive to variations in lattice structure but the latter is.

11.8.2 Mechanism of Anisotropic Etching

Many models have been proposed for the mechanism of silicon anisotropic etching in alkaline solutions. They can be grouped into two categories: those that attribute the relative slower etch rate of (111) planes to the presence of a passive oxide film on the surface, and those that consider the etch rate difference among different orientations to be governed by reaction kinetics.

The models based on surface passivation suggest that a passive layer, similar to the silicon oxide formed under an anodic potential, exists on (111) silicon at OCP but not on other planes [82, 126, 156]. Instead of oxide, formation of inactive hydration complexes of K^+ and OH^- has been proposed to block the (111) surface [49, 134]. The fact that the relative etch rates of the different planes vary with the type of solution was attributed to the orientation-dependent adsorption of solvation complexes on the surface.

The passivation models are not in agreement with a number of experimental observations. Firstly, it has been found that the (111) surface in KOH at OCP is free of oxide and is terminated by hydrogen [50, 157]. Also, the passivation potential for (100) surface in KOH solutions is more negative than that of (111) surface indicating that it is easier to passivate (100) surface than the (111) surface [128, 139]. Furthermore, the etching of silicon at OCP in alkaline solution is mainly a chemical reaction (no involvement of carriers), while the formation of oxide is an electrochemical one [18, 19, 158]. In addition, the etch rate of (111) surface in KOH, although much smaller than those of (100) and (110) planes, have definite values, in the range of 2–10 A s^{-1}, which is much larger than the dissolution rate of passive oxide film in KOH in the order of 0.01 A s^{-1} [126].

The basic consideration in reaction kinetics models is that the reaction rate is determined by the lattice structure on the surface. The difference in the lattice structures of various crystal planes gives rise to differences in surface bond density, electron density, surface free energy, and so on, which then determine the dissolution rate of the surface silicon atoms. The rate of this reaction is determined by the bonding condition of the surface silicon atoms. The difference in etch rate between (111) and (100) surfaces was considered to relate to the bond densities on the two surfaces in the early models [18]. It was soon recognized that this effect alone will not cause an etch rate difference of more than a factor of 2. Some models attributed the anisotropic etching to the different energy levels of surface states on (111) and (100) surfaces arising from binding surface atoms to OH^- [45]. The difference in the activation energy of various crystal planes was related to the difference in the energy states of the backbonds of differently oriented surfaces. Such a difference in energy levels has also been attributed to the difference in work function between (111) and (100) planes [139].

A number of studies attributed lattice defects to be the cause of anisotropic etching. On the basis of the finding by STM imaging, Allongue et al. [19, 157] proposed that the anisotropic etching occurs preferentially at the kink-site atoms at the edges of terraces. Jakob and Chabal [158] reasoned that the silicon

atoms, such as those at the steps and kink sites, that are most physically accessible are preferentially attacked. The reactants and the dissolved complex have a certain physical dimension and orientation so that certain pathways may be forbidden due to steric constraints.

Elwenspoek [159] used the quantity of step free energy, which is the free energy required to generate steps on a flat surface, to measure the effect of surface defects on the anisotropic etching rate. The etching kinetics of a crystal is governed by the kinetics of step formation and removal. The rate-determining step in the dissolution process is the step generation. The etch mechanism of a smooth face is characterized by a nucleation barrier associated with the step free energy. The difference between the etch rates of crystal planes (111), (110), and (100) is thus due to the difference in step free energy, which is the largest for the (111) surface. For a rough face with many steps, for example, a misoriented crystal plane, there is no nucleation barrier as the step free energy is near zero. In such a case, atoms may be removed from the surface without changing the number of steps.

11.8.2.1 A General Model for both HF and KOH Solutions

A coherent mechanistic model has to address three basic aspects: (1) the physical cause for the difference in the removal rates of individual atoms from the surface of different crystal orientations, (2) the kinetic processes that actualize such physical cause, and (3) the surface condition that determines the removal rate of the surface atoms. Also, the model has to be consistent with the data from the alkaline etching and HF etching.

The physical cause in the first aspect can be considered to be, most convincingly, associated with the backbond strength theory [18, 157, 158]. Because of the adsorption of OH^-, which has a large electronegativity, the backbond is polarized and weakened. The {111} plane has only one Si−OH surface bond, whereas {100} surface has two bonds as characteristically represented by the scheme shown in Fig. 28. The atoms on the (100) surface thus have weaker backbonds compared to those on the (111) surface due to the larger polarization. This is considered the fundamental cause for the higher dissolution rate of (100) surface than that of (111) surface.

Regarding the second aspect, generally, as shown in Fig. 29, three species may be involved in the etching reaction: charge carriers, that is, electrons and holes, at the semiconductor surface; chemical species such as OH^-, NO_3^- H_2O, and so on near the surface in the solution; and active surface silicon atoms which are favorable sites for reaction and removal. Unlike the other two species, charge carriers may or may not be involved depending on whether the reaction is of an electrochemical nature. The concentrations of each of these species are determined by different processes such as diffusion, migration,

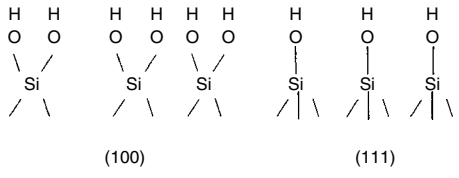


Fig. 28 Schematic illustration of the bonding condition of (100) and (111) surfaces.

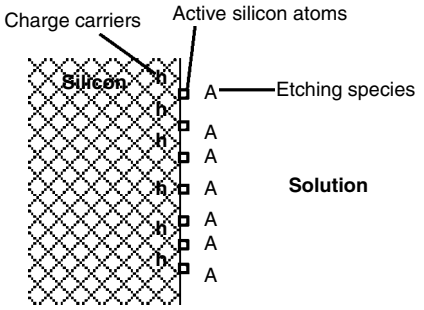


Fig. 29 Schematic illustration of the elementary species involved in the dissolution reaction of silicon; h: carries inside the semiconductor; ◇: active silicon atoms on surface; A: etching species in the solution.

adsorption, solvation, and so on. In particular, the concentration of surface-active silicon atoms depends on crystal orientation, which is the third aspect to be discussed below.

Since the anisotropic etching in alkaline solutions occurs on a bare surface of silicon, formation of an oxide film will mask the crystallographic differences among the surfaces of different orientations and thus result in isotropic etching. Also, experimental results indicate that while isotropic etching may occur under either diffusion control or activation control, anisotropic etching can only occur by processes that are surface controlled at least for the slow etching surfaces. Anisotropic etching does not occur when the etch rates of all the surfaces, including the ones with the slowest etching rates, are controlled by the supply of the solution species. Whether charge carriers are involved in etching reactions is also important. Charge carriers are involved in the isotropic etching in HF solutions but not in the anisotropic etching in KOH solutions.

The third aspect in the etching mechanism is about the nature and concentration of the active surface silicon atoms. Since the stability of surface atom depends on the number of backbonds, it can be proposed that the probability of atom removal from a perfectly (111) surface lattice is very small and etching on (111) surface proceeds only at lattice inhomogeneities such as steps, kinks, and vacancies. Thus, a perfect (111) surface will have an extremely low etch rate and the etch rate of real (111) surfaces is determined by the etch rate at the steps and the density of the steps.

The key points of this third aspect can be summarized as (1) the etch rate of a perfect (111) surface is negligibly small, (2) the etching of real (111) surfaces is due to the etching along the edges of (111) steps, (3) the amount of the (111) steps depends on the degree of misorientation from the perfect (111) plane, (4) all crystallographic planes can be viewed as composed of the (111) terraces and steps at a microscopic level. The etch rate of a surface with an angle from (111) surface can then be correlated with that at (111) steps as shown in Fig. 30. The vertical etch rate, V_v, which is the planar etch rate normally measured in experiments, is

$$V_v = V_s a/b \tag{10}$$

where a is the step height, b the terrace width, and V_s the step etch rate. For a small angle θ, it becomes

$$V_v \approx V_s\theta \tag{11}$$

Thus, the etch rate of a surface is proportional to the angle of deviation from

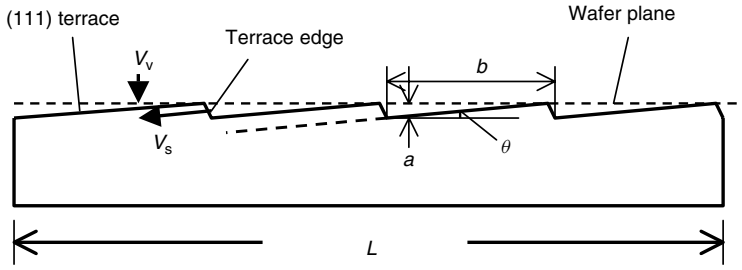


Fig. 30 Relationship between vertical etch rate, V_v, step etch rate, V_s, and the angle from the (111) plane, θ.

(111) surface. For a given surface with a dimension of l the above equation can also be expressed as

$$R_v = R_s an/l \tag{12}$$

where n is the number of steps over the length l. Thus, the larger the misalignment from the (111) surface the more the terraces and steps, and thus the larger the etch rate. This means that the surfaces with large angles to the (111) surface have higher etch rates relative to the (111) surface. Essentially, this implies that at the atomic scale etch occurs only at the steps of the most inert surface, that is, the (111) steps in a direction parallel to the (111) terrace.

In the case of a perfectly oriented (111) surface, such as the side walls of etched cavity on a (100) wafer, there may still be a definite etch rate because steps may be generated from surface defects such as vacancies and dislocations. Thus, for a smooth and perfectly oriented (111) surface, the etch rate is limited by the rate of step generation, that is, a nucleation process.

According to the model illustrated in Fig. 29, the dissolution of a silicon surface involves three essential species: etching species A such as OH^- in the electrolyte, active silicon atoms on the surface □, and charge carriers from the semiconductor e. The etch rate of the surface can then be expressed as

$$R = k[A]^a[\square]^b[h]^c \tag{13}$$

where a, b, c, and k are constants. In KOH where the dissolution is predominantly chemical in nature the etching is independent of $[h]$. Equation (13) is reduced to

$$R = k[A]^a[\square]^b \tag{14}$$

On the other hand, in HF solutions in which the dissolution reaction strongly depends on carrier concentration and the etch rate is described by Eq. (13). The reason that the planar etching rate in HF solutions is isotropic can be attributed to the fact that the etched surface in the exponential region is rough at an atomic scale. Therefore, the densities of active silicon atoms on such a rough surface are similar for different orientations. In the electropolishing region, where the dissolution proceeds through oxide formation and dissolution, the etching is isotropic due to the fact that anodic oxide is amorphous and is identical on the surfaces of different orientations. The etching in this region is intrinsically isotropic.

The reactivity of the surface can be generally described by the density of the active surface atoms as illustrated in Fig. 29. The

density of the surface-active atoms determines the difference in the reactivity of different surfaces. Thus, the etching of silicon in KOH is anisotropic because the difference in the density of the surface-active atoms on (100) and (111) surfaces is large. On the other hand, the etching of silicon in $HF-HNO_3$ is largely isotropic because the density of the surface-active atoms is similar for surfaces of different orientations. This is because in HF solutions the surface has an intrinsic tendency to roughen due to the sensitivity of the electrochemical reactions to surface curvature [12]. The surface which is rough at the atomic scale has no distinct crystallographic character and can be viewed as an amorphous surface as illustrated in Fig. 31a. On the other hand, the surface shown in Fig. 31b, which is less rough and has a clear crystallographic character can exhibit anisotropic behavior when the kinetics is limited by the number of active sites. It can be generalized that when the number of atoms at kink sites, steps, and other defects is close to the density of surface atoms, the surface loses its crystallographic nature. Similarly, the reaction is also essentially isotropic in the case of electropolishing in HF solutions because the dissolution is through the formation and dissolution of silicon oxide that is amorphous in structure. The formation of an amorphous oxide film masks the crystallographic nature of the substrate as shown in Fig. 31c.

11.8.3 Basic Features of Anisotropically Etched Surfaces

Anisotropic dissolution of crystal surfaces results in the formation of surface contour whose geometric features depend on the crystal orientation [81]. During steady-state etching the etched surface profile exhibits a characteristic shape: convex or concave [160]. A convex surface will be bounded by fast etching planes while a concave surface will be bounded by slow etching planes.

Etching of a sphere, which is a convex surface, will result in a polyhedron bounded by faces which exhibit high etch rates, and vertices corresponding to minima in etch rates [160]. Since the fast etching planes vary with solution composition the faces and vertices of etched polyhedron also vary. For example, in KOH solutions the polyhedron has six four-sided <100> vertices and eight six-sided <111> ones, defining 24 {320} fast etching planes [161]. In CsOH solution the polyhedron has eight three-sided <111> vertices and six four-sided <100> giving 12 {110} fast etching planes. A sphere etched in hydrazine–water mixture gives eight curved

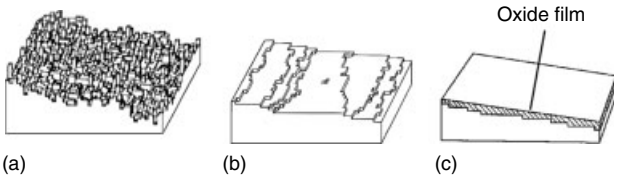


Fig. 31 Schematic illustration of the conditions of surface lattice structure: (a) amorphous-like surface with no identity of orientation, (b) surface with kinks, steps, and terraces characteristic of certain crystalline orientation, and (c) surface with no identity of the lattice structure of the crystal due to the coverage of an amorphous oxide film.

faces with the center of the faces to be (100) planes and the corners point in the <111> directions [91]. In RbOH solution the polyhedron has 14 vertices and 24 curved faces [125].

Using masks of different shapes and orientations and controlling the planes to disappear and develop during etching, various etched features such as cavity or pyramid can be obtained from anisotropic etching [162]. Cavities and mesas etched on silicon in alkaline etchants are bounded by sidewalls, which represent the slowest etching (111) planes of the crystal. The exact geometry of cavities etched through a mask opening is a function of the orientation of the wafer, the geometry of the opening, its alignment relative to wafer's crystal axes, and the duration of etching [130]. Figure 32 illustrates the shape of three cavities anisotropically etched on (100) surfaces through openings of different geometry defined by the masks. After sufficient etching time, the intersection of the cavity and the surface is a rectangle bounded by four (111) planes that enclose the opening defined by the mask.

The final rectangular cavity bounded by four (111) planes independent of the initial opening shape can be explained with the (111) step etching model illustrated in Fig. 30. All the (111) steps that intercept the etching (100) surface and are exposed within the opening defined by the mask are etched in a direction parallel to the (111) planes. The etching along these planes will continue until perfect (111) planes, which define the walls of the etched cavity, are reached and all the exposed (111) steps vanish. Similar to fabrication of a cavity, a convex feature like pyramids can be fabricated by etching a masked (100) silicon surface [68, 163, 164].

For (110) wafer, the geometries of the etched holes are different because the four sets of {111} planes that form the sidewalls of the cavities are not equivalent. Vertical grooves with a high-aspect ratio can be obtained by etching a masked (110) wafer since two sets of {111} planes are vertical to the (110) surface [82].

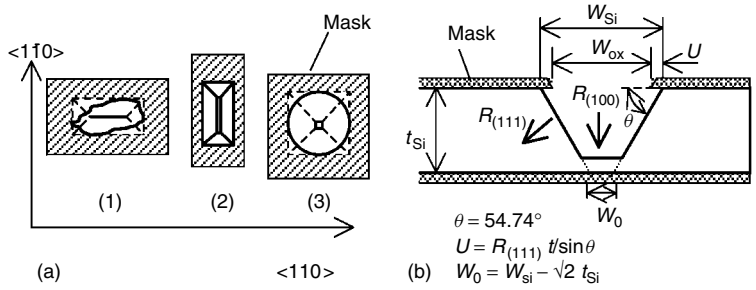


Fig. 32 (a) Anisotropically etched holes through small openings of different geometry in a surface film on a (100) oriented silicon wafer. Four convergent self-limiting {111} etch planes define the sidewalls of each cavity. The geometry of the opening is: (1) arbitrary, (2) rectangular (3) circular. The four sets of {111} planes in a (100) crystal are equivalent and intersect each other at the surface along the [109] directions. (b) Cross section through anisotropically etched pattern in Si(100). The orifice dimension $W_o = W_{Si} - t_{Si}$, and the underetching U are a function of the etch rate $R_{(111)}$ of the {111} planes and etching time t for an accurately aligned pattern and three defect-free Si–SiO_2 interfaces (after Bassous [130]).

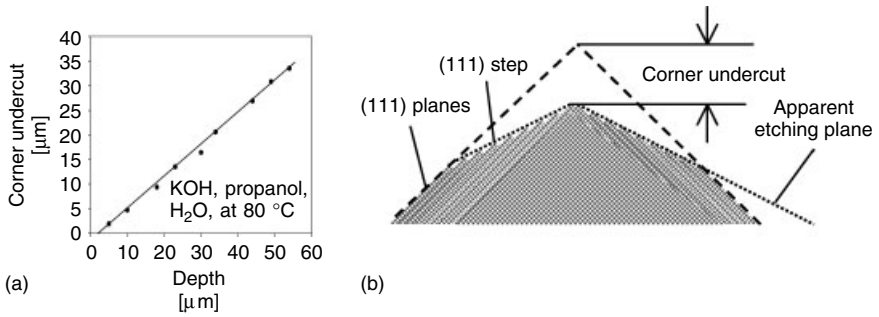


Fig. 33 (a) Amount of undercut as a function of etching depth in KOH/propanol + H_2O at 80 °C (after Bean [131]). (b) Schematic illustration of corner undercutting based on the (111) step etching model shown in Figs 29 and 30.

For a convex square or rectangular structure, etch undercutting occurs at a corner, which can be a problem in the etching for deep structures [101, 131]. The amount of corner undercutting is proportional to the depth of the etching feature as shown in Fig. 33a [131]. Such clear proportionality is due to the definite etch rate of the particularly fast etch planes in the etchant. Because the fast etchant planes are etchant specific, the surfaces that define the etched corner may be different depending on the etchant. The planes are found to be {331} in KOH–propanal–H_2O etchant [118, 131]. They are {411} planes in pure KOH solutions [153, 165] and can vary with KOH concentration [24]. In KOH–propanol–water solutions the etched corner is bounded by {331} planes [131]. In hydrazine–water solutions, {211} planes are found to define the etched corner [91, 166]. The problem of corner undercutting can be circumvented using compensation masks [91, 165].

Etching of a corner defined by (111) planes can be explained by the (111) step etching model illustrated in Fig. 30. The etching of perfectly oriented (111) planes are determined by the generation of steps. The corner, being the terminating line of the two (111) planes, can serve as the site to generate (111) steps. Once a (111) step is generated it will be etched at a rate determined by the etching conditions. The apparent orientations of the two etched surfaces are determined by the rates of step generation and step etching as illustrated in Fig. 33b. A fast step etch and a slow step generation will result in a surface with a small angle to the (111) plane.

11.9 Surface Roughness

Surface roughness, which appears as a series of peaks and valleys on a microscopic scale, is a measure of the degree of unevenness of a solid surface. At an atomic scale, it describes the uneven termination of the lattice at the surface relative to a perfectly flat reference plane. Roughness can be described at two levels: microroughness and macroroughness. The microroughness describes the surface unevenness due to atoms, vacancies, kinks, and steps associated with the few layers of atoms on the surface and macroroughness describes the

surface topological variations beyond the first few atomic layers. Macroroughness usually exceeds the microscopic roughness by several orders of magnitude [167]. A surface that is macroscopically (over a large area) rough may be microscopically (over a small area) smooth.

Surface roughness is an important factor in the performance and reliability of devices having microscopic dimensions [136]. It can essentially affect all aspects of silicon technology [168, 169]. To guarantee the performance of electronic devices the microstructures fabricated from silicon are required to be smooth, with roughness much smaller than the dimensions of the structure components [127]. As the device dimension gets smaller the device will be composed of thinner oxides and shallower junctions and the effect of surface roughness will become critical.

11.9.1 Microroughness

Microroughness, or surface unevenness, at the atomic scale is always present on a real surface whether it is smooth or rough at a macroscopic scale. For a macroscopically smooth surface the microroughness is mainly determined by the wafer polishing and cleaning processes that are used to prepare the surface for further processes. As-received wafers generally have very small microroughness, typically 0.15–0.4 nm (rms) [136, 170, 171].

Any cleaning process can alter microroughness. For example, cleaning with the RCA process is found to increase the microroughness of as-received wafers from 0.2 nm (rms) to about 0.4 nm [136]. Any modification to each of the steps used in the RCA cleaning can also affect the microroughness [87, 111, 136, 158, 170].

Etching in HF-containing solutions is particularly important since they are involved in almost all cleaning processes for silicon surfaces. It is often the last cleaning step in surface preparation for further processing. The silicon surface that is treated with HF solution is terminated by hydrogen. The atomic level flatness is related to the concentrations of monohydride, dihydride, and trihydride [172–174]. On a (111) surface monohydrides tend to form on the perfect lattice while dihydrides and trihydrides tend to form at steps and defects. An atomically flat (111) surface is dominated by monohydrides and a straight step is dominated by dihydrides. Silicon surfaces, either (111) and (100) orientations, after being treated in HF solutions are generally rough at an atomic scale and have surface features such as steps, kinks, and defects [158, 174, 175].

The microroughness of silicon surface in various cleaning solutions generally increases with time, that is, with the amount of material removal [111, 136, 171]. Thus, the small roughness observed on the surface treated in cleaning solutions is generally associated with a small amount of material removed and larger surface roughness with removal of a large quantity of material.

11.9.2 Macroroughness

The macroroughness of a silicon surface treated in a solution can be determined by numerous factors associated with the initial surface condition, the crystal orientation, solution composition, and treatment procedure and time.

In HF solutions, in which silicon etches only very slowly, a silicon surface tends to become rough with lapse of time due to the formation of etch pits

over prolonged etching [171]. Similarly, roughness tends to develop in BHF solutions due to surface pitting [111]. In $HF-HNO_3$ solutions, in which etch rate is high, the roughness of etched surfaces depends on the composition of the solution. For a given HF concentration, there is a concentration range of HNO_3 within which smooth surface can be obtained [76, 176]. In $HF-CrO_3$ solutions the surface roughness is found to decrease with decreasing HF/CrO_3 ratio [62, 63].

In KOH solutions, the roughness varies with concentration with a peak at about 5 M and it is lower in stirred solutions than in nonstirred solutions at room temperature [127]. In nonstirred solutions, the roughness is parallel to the etch rate which also shows a peak at about 5–6 M, but in the stirred solutions it remains at a relatively small value in the whole concentration range. The roughness of (100) etched surface decreases with KOH concentration at different temperatures [78, 169]. Roughness in KOH tends to decrease with increasing temperatures. Cleaning with different water and HF rinses have little effect on the roughness of the etched surfaces in KOH solution, implying that native oxide on the surface is not important in determining the roughness [127, 167, 177]. The composition of solution is also important in affecting the surface roughness in other alkaline solutions [91, 95, 109, 140].

The roughness of the silicon surface changes with many operating conditions. It is a strong function of potential [9, 38, 129, 178]. Addition of parts per million levels of surfactants can significantly change the etch rate and surface roughness in KOH solutions [167] and in fluoride-based solutions [111, 179]. Wafer growth process [169], flow rate etchant [180], and position of wafer in cleaning tank [177] are also found to affect the roughness.

Macroroughness varies with the amount of material removed [127, 152, 169]. For a given etching system, there is a steady-state roughness. If the initial roughness is smaller than this roughness the roughness of the etching surface tends to increase, and if the initial roughness is larger than the steady-state roughness it tends to decrease. For example, in wafer etching, the surface roughness increases with time on the polished front side (initial roughness 2.5 nm) but decreases with time on the backside (initial roughness 300 nm) [167]. Generally, as-received silicon wafers have a very small surface roughness, and in most etchants the roughness tends to increase as etching proceeds. The effect of time depends on orientation; the roughness of the (110) plane has been found to increase with time while those of (111) and (100) planes remain relatively constant [78]. In KOH solutions, it was observed that as soon as the duration of etching exceeds a critical time, about 15 minutes, characteristic pits or hillocks develop on the etching surface [181].

11.9.3 Crystallographic Characters and Formation of Hillocks

Surface roughness of silicon crystals has clear crystallographic characteristics as illustrated in Fig. 34 [78]. On a microscopic scale, roughness is associated with lattice steps, vacancies, and so on, which are determined by the lattice structure of the surface. At a macroscopic level, crystallographic character may be revealed in the topographic features, for example, the hillocks formed on (100) surface.

The macroroughness of etched silicon surface in alkaline solutions is a function

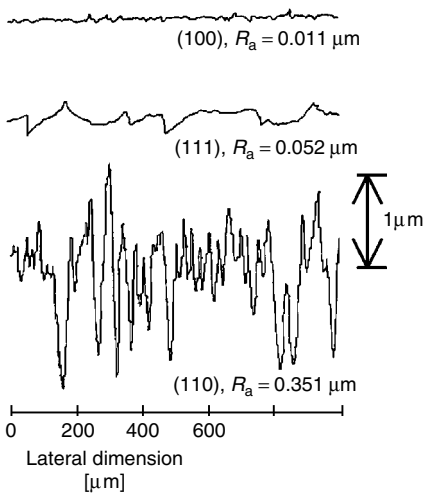


Fig. 34 Surface roughness of three principal Si crystal orientations after etching for 10 minutes (after Sato et al. [78]).

of surface orientation [45, 81, 181]. For example, Sato et al. [78] found that for etching KOH the smoothest surface appears at the (100) plane while the (320) and (210) planes, which also have higher etch rates than other planes, are very rough.

The (100) surface tends to roughen quicker than (111) surface and the roughness tends to be permanent on (100) surface while it tends to be transient on (111) surface [47]. Such crystal orientation-dependent roughness could be explained by the anisotropic etching model illustrated in Fig. 30. The preferential etching at the (111) steps of the (111) terraces results in the removal of the terraces and reduction of the (111) steps and a reduction of microroughness.

In particular, the hillocks formed on (100) surface are crystallographic structures bounded by four (111) planes resulting from the anisotropic etching and their formation directly contributes to the roughness of the surfaces [127]. Hillocks may, under certain conditions, form in HF-based solutions [158], but most commonly in alkaline etchants [88, 109, 127]. Since (111) surface etches very slowly the formation of hillocks leads to the decrease of etch rate [88, 127].

The extent of hillock formation in alkaline solutions depends strongly on solution composition and operating conditions. Solution stirring and higher solution temperature reduce the hillock-associated surface roughness in KOH solutions [45, 127]. In NH_4OH solutions it occurs for all concentrations, temperatures, stirring conditions, and sample preparations [86, 88]. Hillocks tend to readily form on poorly cleaned surfaces and in etchants containing excessive amounts of dissolved silicon [109, 182]. They may result from local masking by contaminants or deposits of reaction products at the apex of the pyramidal hillocks. The anisotropic dissolution of the material around the masked area results in the formation of pyramid. The masking substances can be surface contaminants, precipitates in silicon material, surface deposits, and formation of gas bubbles [109, 127].

Formation of hillocks can be reduced by controlling solution composition,

temperature and agitation, electrode potential, surfactants, pressurizing, and so on [88, 95, 127]. In NH_4OH solutions it can be suppressed by the addition of H_2O_2. In TMAH solutions at concentrations lower that 25% small hillocks start to develop and the size and number increase with decreasing concentrations [78, 95]. Addition of IPA in TMAH increases the size of hillocks but does not change hillock density. Electrode potential can also affect hillock formation [129]. Increasing the pressure above the etchant bath is found to enhance the formation of hillocks in KOH solutions [121].

11.9.4 Origins of Roughness

Roughness developed in a cleaning or etching solution is a result of uneven dissolution across the crystal surface. Many factors in an etching process may contribute to the uneven distribution of the dissolution rate at both micro- and macroscales. In general, any process that causes temporary or permanent surface inhomogeneity will result in preferential dissolution of some areas relative to other areas.

The causes for roughness may be grouped into two categories, process dependent and material dependent as shown in Fig. 35 [3, 127]. The process-dependent causes are the deposition of metal or nonmetal particles (Fig. 35a,b) or bubbles (Fig. 35c) which provide masking to the deposited areas. The difference in the masking effect between a metal and a nonmetal particle is that metal can actively participate in the dissolution reaction by a galvanic action. The material-dependent causes, such as preferential dissolution at lattice defects (Fig. 35d), anisotropic dissolution associated with terraces and steps (Fig. 35e), and formation of pores (Fig. 35f) are associated with the structure and property of silicon crystal.

The surface of silicon crystal, no matter how it is finished, will have a certain number of lattice defects, which tend to dissolve preferentially, resulting in formation of etch pits and other features. Also, a silicon surface, whether initially smooth or not, in HF solutions, has an intrinsic tendency to roughen and form micropores governed by sensitivity of the electrochemical reactions on a semiconductor electrode to surface curvature [12]. Furthermore, the two groups of factors shown in Fig. 35 may affect each other. The initial lattice

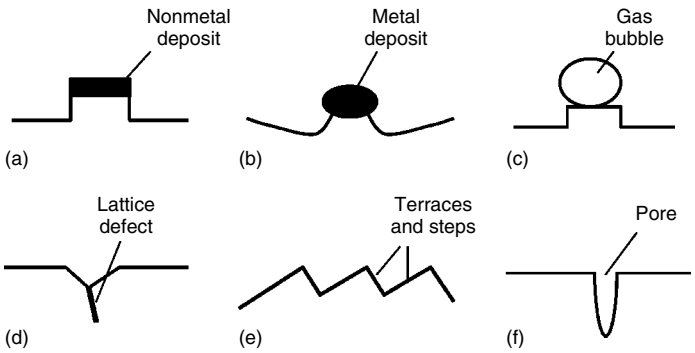


Fig. 35 Schematic illustrations of the possible origins of surface roughness.

inhomogeneities may provide the sites for deposition while localized deposition may enhance the development of etch features such as pits or hillocks.

11.10 Applications of Etching

The applications of etching can be broadly classified into three categories: (1) surface preparation, (2) structural characterization, and (3) device fabrication [160]. In category (1) etching is used to prepare suitable surfaces for subsequent physical and chemical measurements or processing. Here etching is used as a tool for cleaning or polishing, for removing the damaged or defective surface layer, or for surface stabilization. In category (2) etching is used for the identification of defects such as dislocations, for microstructural studies, for orientation examination, or for impurity distribution. In category (3) etching is used for removing materials and for machining structures in device fabrication.

11.10.1 Cleaning

A clean surface is essential for device reliability and performance [183]. It becomes critical as the dimensions of devices become smaller and smaller as a result of ever-increasing integration and complexity. It has been estimated that over 50% of yield losses in integrated circuit (IC) fabrication are due to microcontamination [184]. Today, a typical process flow for advanced ICs consists of 300 to 500 steps, 30% of which are wafer cleaning steps [185]. The need for wafer cleaning can be separated into three areas (1) preparation of the wafer surfaces for oxidation, diffusion, deposition, and metallization, (2) preparation for the application of photoresist, and (3) removal of photoresist after the etching process [186, 187].

A cleaned silicon wafer is (1) particle free, (2) without organic contaminants, (3) with minimal metallic contamination, (4) free of native oxide, and (5) with minimal surface roughness [24, 58]. A clean surface has a surface concentration of contaminants such as Fe, Cu, and Ni of less than 10^{10} cm^{-2} and microroughness of about 0.2 nm [136]. As-received wafer typically has several thousand particles, a roughness of about 0.2 nm and metal concentrations of about 5×10^{10}–10^{12} cm^{-2} for Fe, Cu, and Zn [135, 188, 189]. The degree of cleanness is constantly improved as cleaning technology continues to progress.

Contaminants on silicon surfaces can be classified as molecular, ionic, and atomic or as hydrocarbons, metals, and particles [57, 190]. Typical molecular contaminants are waxes, resins, oils, and organic compounds that are commonly generated in the processes after sawing, or from human skin and plastic containers. They are usually attached to the surface by weak electrostatic forces. Ionic contaminants such as Na^+, Cl^-, F^-, and I^- are present after etching in HF or alkaline etchants. They may be precipitated on the surface by physical adsorption or chemical adsorption. The atomic contaminants of concern are metals such as gold, iron, copper, and nickel originating from acid etchants. The metallic impurities can only be effectively removed by wet cleaning processes [58, 136]. The ceramic particles such as SiC, Si_3N_4, Al_2O_3, SiO_2, and C, originate to a large extent from wet processes and cleaning baths. These particles are attracted by the electrical field generated by the wafers [190]. The contaminants may physically locate in different phases at the

silicon/electrolyte interface depending on whether the surface is covered with an oxide film and whether the contaminants are adsorbed onto the surface before, during, or after the formation of the oxide film.

11.10.1.1 RCA Cleaning

The RCA process, which was originally developed by Kern and Puotinen in 1970 [57], is still the basis of current wet cleaning process for bare or oxide covered silicon surfaces. The original RCA cleaning, which tends to etch 30–40 Å silicon substrate during the process [55], is based on two hydrogen peroxide solutions: an alkaline mixture with ammonium hydroxide followed by an acidic mixture with HCl [184].

In the first treatment, a hot solution of NH_4OH (27%), H_2O_2 (30%), and H_2O, with a proportion of 1:1:5 to 1:2:7 by volume of the chemical agents, known as the *SC1 cleaning solution*, is used [184]. The SC1 solution is designed to remove organic contaminants by both solvating action of the ammonium hydroxide and the oxidizing action of the peroxide. The oxidizing action also results in the formation of surface silicon oxide, which dissolves in the solution. It is thus very effective in particle removal as continuous oxidation and etching of the oxide take place under the particles [170]. The metal surface contaminants such as gold, silver, copper, nickel, cadmium, zinc, cobalt, and chromium are also oxidized by H_2O_2 and dissolved by the complexing effectiveness of the ammonium hydroxide [136, 184, 189].

The second treatment step exposes the rinsed wafer to a solution known as the *SC2 solution* which is a hot mixture of H_2O_2 (30%), HCl (37%), and H_2O in the proportions of 1:1:6 to 2:1:8 by volume. This cleaning solution is designed to remove alkali ions and cations such as Al^{3+}, Fe^{3+}, and Mg^{2+}. It also cleans metallic contaminants that are not entirely removed in SC1 solution, such as gold [57]. A preliminary cleaning step with a hot H_2SO_4–H_2O_2 mixture (2:1) can be used for grossly contaminated wafers having visible residues. The SC1 solutions are continually modified for better results, dilution [55], ultrapure water [190], ozonized water [58] and other additions [44, 136, 184]. More recently, this cleaning solution is often used with a concentration only one-tenth of the original solution to avoid surface roughness and to reduce production cost and its effect on the environment [191].

Cleaning of metallic impurities on a silicon surface requires the oxidation and dissolution of the metal atoms. Most metals have a standard potential that is much higher than that of silicon. To oxidize the metal atoms, the electrode potential, at least in the local area of the deposits, has to be raised to the level of the reversible potential of the metal. In the SC1 cleaning solutions the potential of silicon is 0.4 V_{sce} at which the silicon surface is passivated (It is −1.10 V_{sce} without H_2O_2) [84]. This means in the SC1 solution the species with reversible potential less than 0.4 V_{sce} will be oxidized. Passivation, which is responsible for the slow etch rate in the SC1 solution [55], serves three purposes: (1) it allows the surface to be anodically polarized, which is required for oxidizing of metal impurities, (2) it prevents much etching of the substrate during the cleaning process, and (3) it helps to physically dislodge the metal atoms from the surface by forming a oxide layer between the metal atoms and the silicon surface.

11.10.2 Defect Etching

Defect etching refers to the etching process that preferentially attacks the strained

bonds of defects within a crystal. It is a simple and fast method to determine the structural perfection of a single crystal. The morphological features developed by defect etching have a characteristic shape and may occur at random or in arrays revealing crystalline imperfections such as dislocation, slip, lineage, and stacking faults. Defect etching has been widely used to evaluate dislocations and also to delineate process-induced defects in silicon wafers and epitaxial deposits [192, 193].

Many different etchants have been developed for the evaluation of different types of crystal defects such as: flow-pattern defects [115], stacking faults [74, 192, 193], dislocations [69, 72–74, 194], dislocation network [72, 193, 195], oxide precipitates [196], swirl patterns [74, 75, 148], striations [74, 197], hillock defects [198], epitaxial defects [192], epitaxial alignment [199], grain boundary [69, 72, 200], twin band [69], diamond saw damage [201], *pn* junction [202–204], metallic precipitates [205], and damaged layer of mechanically polished surface [206].

Most defect etchants are mixtures of CrO_3 and HF. These etchants are easy to use and fast, taking several minutes to reveal the defect etch features. Non-CrO_3 etchants are also used for defect etching. For example, $HF-HNO_3-CH_3CO_2H$ based etchant has been found to be highly sensitive for defects of polycrystalline materials [69, 207]. Etching–staining in a $CuSO_4-HF$ solution reveals surface defects associated with dislocation networks [195], and *pn* junction [204, 208]. The copper that is preferentially deposited in the vicinity of the etched area provides feature delineation.

Etching of dislocations is most pronounced on (111) among the major surfaces due to the large difference in surface potential between a dislocation and the surrounding area [209]. The etch pattern of dislocations is determined by the inclination of dislocations to the surface [75]. The basic unit of an etch pit is generally bounded by the (111) planes intersecting to the surface.

11.10.3 Material Removal

Etching is widely used as a process for material removal in silicon device fabrications. The important aspect of any etching process is the ability to control the amount of materials with sufficient spatial accuracy. The process and structural diversity in device fabrication necessitates diverse etching techniques, chemical or electrochemical, anisotropicity, masking, and illumination to provide uniform as well as selective etching. Figure 36 illustrates the basic uniform and selective etching techniques for removing materials.

11.10.3.1 Uniform Material Removal

In the category of uniform etching the simplest method is immersing silicon samples in etchants such as KOH or $HF-HNO_3$ solutions (Fig. 36a). It is normally used in applications such as damaged layer removal [152, 210] and polishing [38, 211], and in-depth profiling [212]. The etching is most commonly operated at the OCP but can be done under an anodic potential and/or illumination to add extra control. Etching with potential control requires an ohmic contact at the backside of the wafer to apply an anodic potential (Fig. 36c). Also, for uniform etching in fluoride-based solutions, the potential has to be in the current plateau region of an anodic $i-V$ curve (see Fig. 1). Instead of

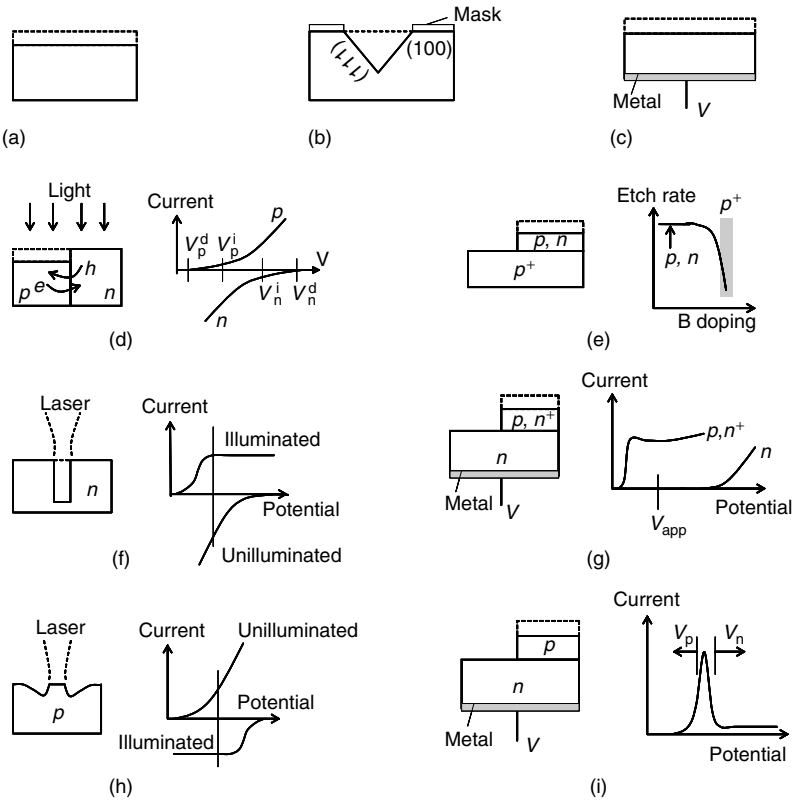


Fig. 36 Schematic illustration of the different etching techniques for material removal : (a) planar etching at OCP, (b) anisotropic etching of masked surface, (c) planar etching under an anodic bias, (d) lateral selective etching of the *p* region of an illuminated *pn* junction, (e) vertical selective etching using heavily doped material for an etch stop, (f) preferential etching of illuminated area on an *n*-type material, (g) vertical selective etching of *p* or n^+ material under an anodic potential in the dark using the low etch rate of *n*-Si as an etch stop, (h) preferential etching of the nonilluminated area on an illuminated *p*-type material, and (i) vertical selective etching of *p*-Si of anodically biased *pn* junction using the passivity of *n*-Si as an etch stop.

applying an external current, anodic potential can also be applied through galvanic action by depositing a metal on a part of the silicon surface [213].

Etching of the silicon wafer in mixtures of HF and HNO_3 is routinely used to remove surface work damage and stress from the slicing operation as well as to provide a relatively nonporous and clean backside for the finished wafer [214]. KOH can also be used for removal of a damaged layer [152].

Chemomechanical polishing, which combines mechanical abrasion and chemical etching, is a basic processing technology for the production of flat, defect-free reflective surfaces [215]. The polishing proceeds by a combination of

the reactions of silicon with the solution and the mechanical removal of the reaction products by the polishing pad and the abrasive fluid. Electropolishing has also been proposed for wafer polishing [38, 211, 216].

11.10.3.2 Selective Material Removal

Many different structures can be micromachined on silicon through etching. While IC technology usually employs planar processes, in micromechanics, sensors, and actuators selective etching has to be used to fabricate three-dimensional and mobile elements [217, 218]. Two aspects are of particular importance in the fabrication of structures by etching: etch-stop and etch-feature definition, both relying on the selective etching of materials on a patterned silicon surface [2, 150]. Virtually all microstructures utilize at least one etch-stop technique during the course of their fabrication. Etch-feature definition, on the other hand, is mainly achieved through masking and anisotropic etching [91, 218].

The many etching characteristics of silicon and the numerous etching systems provide a large range of variation in the etch rate. This range of etch rates, in combination with various etching techniques, provides many methods in selective removal of materials on silicon as illustrated in Fig. 36, allowing the fabrication of diverse structures on silicon.

Selective etching can be realized by using a mask, anisotropicity, *pn* junction, focused illumination, heavy boron doping, and potential control. Etching over a confined area on (100) substrate in an anisotropic solution results in the formation of a cavity as illustrated in Fig. 36b. In the case of an illuminated *pn* junction as shown in Fig.36d, the photogenerated holes in the *n* region are flown to the *p* region resulting in dissolution of the silicon in the *p* region. When the illumination is confined in an area, the photo-enhanced dissolution occurs either within the illuminated area or outside the illuminated area depending on the type of material [219–221]. For *n*-type materials, illumination generates a cavity at the illuminated area as shown in Fig. 36f. On the other hand, for *p*-type material, illumination results in the formation of a column as shown in Fig. 36h.

In-depth selective etching requires that the material to be removed has a much larger etch rate than the material beneath it so that etch stop occurs at the end of the etching. For etching in alkaline solutions, heavily boron doped silicon can be used as the material for etch stop (Fig. 36e). This is based on the difference of several orders of magnitude in the etch rates between heavily doped and lowly doped materials. Etch stop can also be realized using the difference in the dissolution rates of *p*- and *n*-type silicon at anodic potentials in HF solutions (Fig. 36g).

In-depth selective etching of silicon in alkaline solutions can also utilize the difference in the passivation potentials between *p*- and *n*-type materials [135, 138, 180, 222, 223]. In this method, as shown in Fig. 36i, an anodic voltage sufficient to cause passivation of *n*-Si is applied through an ohmic contact. Because of the potential drop in the reversely biased *pn* junction, the potential of the *p*-Si maintains at a potential negative to the passivation potential and is etched. Upon complete removal of the *p*-Si, the junction disappears and the etch stops because the *n*-Si is passivated. A current peak, corresponding to the formation of the passive oxide film on the *n*-Si, occurs when the *p*–*n* interface is reached. This current peak signals the onset of the etch stop. Extra control can be obtained by applying

different potentials on the *p*-Si and *n*-Si regions with *p*-Si maintained at a value near OCP and *n*-Si at a value positive of the passivation potential [223]. This prevents the unwanted passivation of the *p*-Si, which can result from short circuit due to point defects and current leakage. Etching of *p*-Si can also be realized by selective passivation using an accumulation layer instead of a *pn* junction for etch stop [224]. Similarly, laterally selective etching can be obtained by biasing a laterally positioned *pn* junction [135].

Another group of methods for selective etching utilizes the high reactivity of PS [3, 225]. PS can selectively be formed in a patterned area. The PS can then be easily dissolved in a KOH solution leading to the selective dissolution of the patterned area. Also, straight-hole array with high-aspect ratios can be generated through the formation of macropores by back illumination on *n*-Si [7, 226].

A locally confined etchant has been used to preferentially etch small holes on silicon [225, 227, 228]. In this technique, an active etchant is generated through a reduction reaction at the tip of a fine electrode that is positioned near the silicon surface. In situations where a diffusion-controlled process limits the in-depth etch rate, deep etching can be obtained using a centrifugal force [229].

References

1. H. C. Gatos, M. C. Lavine, in *Progress in Semiconductors* (Eds.: A. F. Gibon, R. E. Burgess), CRC Press, Boca Raton, FL, 1965, Vol. 9.
2. K. E. Petersen, *Proc. IEEE* **1982**, *70*(5), 420.
3. X. G. Zhang, *Electrochemistry of Silicon and its Oxides*, Kluwer/Plenum Publishers, New York, 2001.
4. C. M. Hu, *Nanotechnology* **1999**, *10*, 113.
5. M. Jacoly, *Chem. Eng. News July*, **2000**, 32.
6. R. F. Service, *Science* **1999**, *284*, 578.
7. V. Lehmann, H. Foll, *J. Electrochem. Soc.* **1988**, *135*, 2831.
8. J.-N. Chazalviel, M. Etman, F. Ozanam, *J. Electroanal. Chem.* **1991**, *297*, 533.
9. X. G. Zhang, S. D. Collins, R. L. Smith, *J. Electrochem. Soc.* **1989**, *136*, 1561.
10. R. L. Meek, *J. Electrochem. Soc.* **1971**, *118*, 437.
11. R. Memming, G. Schwandt, *Surf. Sci.* **1966**, *4*, 109.
12. X. G. Zhang, *J. Electrochem. Soc.* **1991**, *138*, 3750.
13. V. Lehmann, *J. Electroanal. Chem.* **1993**, *140*, 2836.
14. M. J. J. Theunissen, *J. Electrochem. Soc.* **1972**, *119*, 351.
15. Y. Kang, J. Jorne, *Electrochim. Acta* **1998**, *43*(16–17), 2389.
16. L. C. Chen, M. Chen, C. Lien et al., *J. Elechochem. Soc.* **1995**, *142*, 170.
17. R. M. Hurd, N. Hackerman, *Electrochim. Acta* **1964**, *9*, 1633.
18. O. J. Glembocki, E. Stahlbush, M. Tomkiewicz, *J. Electrochem. Soc.* **1985**, *132*, 145.
19. P. Allongue, V. Costa-Kieling, H. Gerischer, *J. Electrochem. Soc.* **1993**, *140*, 1018.
20. E. D. Palik, O. J. Glembocki, I. Heard Jr., *J. Electrochem. Soc.* **1987**, *134*, 404.
21. R. N. Dominey, N. S. Lewis, J. A. Bruce et al., *J. Am. Chem. Soc.* **1982**, *104*(2), 467.
22. M. Candea, M. Kastner, R. Goodman et al., *J. Appl. Phys.* **1976**, *47*(6), 2724.
23. V. Bertagna, C. Plougonven, F. Rouelle et al., *J. Electrochem. Soc.* **1995**, *142*, 3532.
24. V. Bertagna, F. Rouelle, G. Revel et al., *J. Electrochem. Soc.* **1997**, *144*, 4175.
25. M. Szklarczyk, J. Sobkowski, J. Pacocha, *J. Electroanal. Chem.* **1986**, *215*, 307.
26. M. Szklarczyk, J. O. M. Bockris, *J. Phys. Chem.* **1984**, *88*, 1808.
27. A. B. Bocarsly, E. G. Walton, M. G. Bradley et al., *J. Electroanal. Chem.* **1979**, *100*, 283.
28. G. Oskam, J. G. Long, A. Natarajan et al., *J. Phys. D: Appl. Phys.* **1998**, *31*, 1927.
29. Y. Nakato, S. Tonomura, H. Tsubomura, *Ber. Bunsen-Ges. Phys. Chem.* **1976**, *80*(12), 1289.
30. D. J. Blackwood, A. Borazio, R. Greef et al., *Electrochim. Acta* **1992**, *37*(5), 889.
31. A. A. Yaron, S. Bastide, J. L. Maurice et al., *J. Lumin.* **1993**, *57*, 67.
32. M. J. Eddowes, *J. Electroanal. Chem.* **1990**, *280*, 297.

33. H. Seidel, L. Csepregi, A. Heuberger et al., *J. Electrochem. Soc.* **1990**, *137*, 3626.
34. M. Matsumura, S. R. Morrison, *J. Electroanal. Chem.* **1983**, *147*, 157.
35. J. Stumper, L. M. Peter, *J. Electroanal. Chem.* **1991**, *309*, 325.
36. H. Gerischer, M. Lubke, *Ber. Bunsen-Ges. Phys. Chem.* **1987**, *91*, 394.
37. S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 1981.
38. D. R. Turner, *J. Electrochem. Soc.* **1958**, *105*, 402.
39. E. A. Ponomarev, C. Levy-Clement, *Electrochem. Solid-State Lett.* **1998**, *1*, 42.
40. V. Lehmann, U. Gosele, *Appl. Phys. Lett.* **1991**, *58*(8), 856.
41. P. Allongue, V. Kieling, H. Gerischer, *Electrochim. Acta* **1995**, *40*, 1353.
42. L. M. Peter, A. M. Borazio, H. J. Lewerenz et al., *J. Electroanal. Chem.* **1990**, *290*, 229.
43. L. M. Peter, J. Li, R. Peat et al., *Electrochim. Acta* **1990**, *35*(10), 1657.
44. E. S. Kooij, D. Vanmaekelbergh, *J. Electrochem. Soc.* **1997**, *144*, 1296.
45. H. Seidel, L. Csepregi, A. Heuberger et al., *J. Electrochem. Soc.* **1990**, *137*, 3612.
46. N. F. Raley, Y. Sugiyama, T. V. Duzer, *J. Electrochem. Soc.* **1984**, *131*, 161.
47. E. D. Palik, V. M. Bermudez, O. J. Glembocki, *J. Electrochem. Soc.* **1985**, *143*, 871.
48. E. D. Palik, H. F. Gray, P. B. Klein, *J. Electrochem. Soc.* **1983**, *130*, 956.
49. O. J. Glembocki, E. D. Palik, G. R. de Guel et al., *J. Electrochem. Soc.* **1991**, *138*, 1055.
50. J. Rappich, H. J. Lewerenz, H. Gerischer, *J. Electrochem. Soc.* **1993**, *140*, L187.
51. S. M. Hu, D. R. Kerr, *J. Electrochem. Soc.* **1967**, *114*, 414.
52. G. Willeke, K. Kellermann, *Semicond. Sci. Technol.* **1996**, *11*, 415.
53. M. Matsumura, H. Fukidome, *J. Electrochem. Soc.* **1996**, *143*, 2683.
54. K. Sato, M. Shikida, T. Yamashiro et al., *Sens. Actuators* **1999**, *73*, 131.
55. T. H. Park, Y. S. Ko, T. E. Shim et al., *J. Electrochem. Soc.* **1995**, *142*, 571.
56. S. B. Felch, J. S. Sonico, *Solid State Technol. September*, **1986**, 47.
57. W. Kern, D. A. Puotinen, *RCA Rev. June*, **1970**, 186.
58. T. Ohmi, *J. Electrochem. Soc.* **1996**, *143*, 2957.
59. J. T. L. Thong, W. K. Choi, C. W. Chong, *Sens. Actuators, A* **1997**, *63*, 243.
60. B. Schwartz, H. Bobbins, *J. Electrochem. Soc.* **1961**, *108*, 365.
61. D. G. Schimmel, M. J. Elkind, *J. Electrochem. Soc.* **1978**, *125*, 152.
62. J. E. A. M. van den Meerakker, J. H. C. van Vegchel, *J. Electrochem. Soc.* **1989**, *136*, 1949.
63. J. E. A. M. van den Meerakker, J. H. C. van Vegchel, *J. Electrochem. Soc.* **1989**, *136*, 1954.
64. P. M. M. C. Bressers, M. Plakman, J. J. Kelly, *J. Electroanal. Chem.* **1996**, *406*, 131.
65. W. K. Zwicker, S. K. Kurtz, *Semiconductor Silicon* (Eds.: H. R. Huff, R. R. Burgess), Electrochemical Society, Pennington, NJ, 1973, pp. 315.
66. No reference.
67. K. E. Bean, *IEEE Transaction on Electron Devices* **1978**, *ED-25*, 1185.
68. N. A. Cade, R. A. Lee, C. Patel, *IEEE Trans. Electron Devices* **1989**, *36*(11), 2709.
69. Y. Kashiwagi, R. Shimokawa, M. Yamanaka, *J. Electrochem. Soc.* **1996**, *143*, 4073.
70. H. Ochs, D. Bublak, U. Wild et al., *Appl. Surf. Sci.* **1998**, *133*, 73.
71. H. Muraoka, T. Ohhashi, Y. Sumitomo *Semiconductor Silicon* (Eds.: H. R. Huff, R. R. Burgess), Electrochemical Society, Pennington, NJ, 1973, pp. 327.
72. F. S. Aragona, *J. Electrochem. Soc.* **1972**, *119*, 948.
73. D. G. Schimmel, *J. Electrochem. Soc.* **1979**, *126*, 479.
74. M. W. Jenkins, *J. Electrochem. Soc.* **1977**, *124*, 757.
75. K. H. Yang, *J. Electrochem. Soc.* **1984**, *131*, 1140.
76. T. C. Chandler, *J. Electrochem. Soc.* **1990**, *137*, 944.
77. F. Ozanam, C. da Fonseca, A. V. Rao et al., *Appl. Spectrosc.* **1997**, *51*(4), 519.
78. K. Sato, M. Shikida, T. Yamashiro et al., *Sens. Actuators* **1999**, *73*, 122.
79. J. B. Price, *Semiconductor Silicon* (Eds.: H. R. Huff, R. R. Burgess), Electrochemical Society, Pennington, NJ, 1973, pp. 339.
80. E. D. Palik, J. W. Faust Jr., H. F. Gray et al., *J. Electrochem. Soc.* **1982**, *129*, 2052.
81. C. R. Tellier, A. Brahim-Bounab, *J. Mater. Sci.* **1994**, *29*, 5953.
82. D. L. Kendall, *Annu. Rev. Mater. Sci.* **1979**, *9*, 373.
83. T. Wang, S. Surve, P. J. Hesketh, *J. Electrochem. Soc.* **1994**, *141*, 2493.

84. J. E. A. van den Meerakker, M. H. M. van der Straaten, *J. Electrochem. Soc.* **1990**, *137*, 1239.
85. F. Wang, Y. Shi, J. Liu et al., *J. Electrochem. Soc.* **1997**, *144*, L37.
86. U. Schnakenberg, W. Benecke, B. Lochel, *Sens. Actuators* **1990**, *A21–A23*, 1031.
87. H. Kobayashi, J. Ryuta, T. Shingouyi et al., *Jpn. J. Appl. Phys.* **1993**, *32*, L45.
88. U. Schnakenberg, W. Benecke, B. Lochel et al., *Sens. Actuators, A* **1991**, *25–27*, 1.
89. K. T. Lee, S. Raghvan, *Electrochem. Solid-State Lett.* **1999**, *2*, 172.
90. M. Mehregany, S. D. Senturia, *Sens. Actuators* **1988**, *13*, 375.
91. D. B. Lee, *J. Appl. Phys.* **1969**, *40*(11) 4569.
92. R. M. Finne, D. L. Klein, *J. Electrochem. Soc.* **1967**, *114*, 965.
93. A. Reisman, M. Berkenblit, S. A. Chan et al., *J. Electrochem. Soc.* **1979**, *126*, 1406.
94. H. G. Linde, L. W. Austin, *Sens. Actuators, A* **1995**, *49*, 181.
95. O. Tabata, R. Asahi, H. Funabashi et al., *Sens. Actuators* **1992**, *A34*, 51.
96. A. Merlos, M. Acero, M. H. Bao et al., *Sens. Actuators* **1993**, *37–38*, 737.
97. B. Schwartz, H. Bobbins, *J. Electrochem Soc.* **1961**, *108*, 365.
98. A. F. Bogenschutz, W. Krusemark, K. H. Locheer et al., *J. Electrochem. Soc.* **1967**, *114*, 970.
99. R. B. Heimann, *J. Mater. Sci.* **1984**, *19*, 1314.
100. W. Van Gelder, V. E. Hauser, *J. Electrochem. Soc.* **1967**, *114*, 869.
101. P. M. Zavracky, T. Earles, N. L. Pokrovskiy et al., *J. Electrochem. Soc.* **1994**, *141*, 3182.
102. E. Herr, H. Baltes, *Sens. Actuators* **1992**, *31*, 283.
103. E. Steinsland, M. Nese, A. Hanneborg et al., *Sens. Actuators, A* **1996**, *54*, 728.
104. D. J. Godbey, A. H. Krist, K. D. Hobart et al., *J. Electrochem. Soc.* **1992**, *139*, 2943.
105. M. Takigawa, T. Satoh, K. Shono, *J. Electrochem. Soc.* **1975**, *122*, 824.
106. R. T. Howe, *J. Vac. Sci. Technol.* **1988**, *B 6*(6), 1809.
107. H. P. Kattelus, *J. Electrochem. Soc.* **1997**, *144*, 3188.
108. H. Linde, C. Whiting, D. Benoit, *Sens. Actuators, A* **1997**, *63*, 251.
109. O. Tabata, *Sens. Actuators, A* **1996**, *53*, 335.
110. H. Fukidome, T. Ohno, M. Matsumura, *J. Electrochem. Soc.* **1997**, *144*, 679.
111. M. Myamoto, N. Kita, S. Ishida et al., *J. Electrochem. Soc.* **1994**, *141*, 2899.
112. K. M. Sancier, V. Kapur, *J. Electrochem. Soc.* **1980**, *127*, 1848.
113. L. Liou, W. G. Spitzer, S. Prussin, *J. Electrochem. Soc.* **1984**, *131*, 672.
114. E. Propst, P. A. Kohl, *J. Electrochem. Soc.* **1993**, *140*(5), L78.
115. W. Wijaranakula, *J. Electrochem. Soc.* **1994**, *141*, 3273.
116. T. L. Chu, J. R. Gavaler, *Electrochim. Acta* **1965**, *10*, 1141.
117. S. K. Ghandhi, *VLSI Fabrication Principles*, John Willey & Sons, New York, 1983.
118. E. S. Kooij, K. Butter, J. J. Kelly, *Electrochem. Solid-State Lett.* **1999**, *2*, 178.
119. H. Robbins, B. Schwartz, *J. Electrochem. Soc.* **1960**, *107*, 108.
120. D. R. Turner, *J. Electrochem. Soc.* **1960**, *107*, 810.
121. A. P. Abbott, S. A. Campbell, J. Satherley et al., *J. Electroanal. Chem.* **1993**, *348*, 473.
122. E. S. Kooij, K. Butter, J. J. Kelly, *J. Electrochem. Soc.* **1998**, *145*(4), 1232.
123. K. E. Petersen, *Proc. IEEE* **1982**, *70*(5), 420.
124. H. Line, L. Austin, *J. Electrochem. Soc.* **1992**, *139*, 1170.
125. K. B. Sundaram, H. W. Chang, *J. Electrochem. Soc.* **1993**, *140*, 1592.
126. J. W. Faust, E. D. Palik, *J. Electrochem. Soc.* **1983**, *130*, 1413.
127. E. D. Palik, O. J. Glembocki, I. Heard Jr., et al., *J. Appl. Phys.* **1991**, *70*(6), 3291.
128. R. L. Smith, B. Kloeck, N. DeRooij et al., *J. Electroanal. Chem. Interfacial Electrochem.* **1987**, *238*, 103.
129. P. M. M. C. Bressers, S. A. S. P. Pagano, J. J. Kelly, *J. Electroanal. Chem.* **1995**, *391*, 159.
130. E. Bassous, *IEEE Trans. Electron Devices* **1978**, *ED-25*, 1178.
131. K. E. Bean, *IEEE Trans. Electron Devices* **1978**, *ED-25*, 1185.
132. J. Nijdam, E. Van Veenendaal, J. G. E. Gardeniers et al., *J. Electrochem. Soc.* **2000**, *147*, 2195.
133. P. Akhter, A. Baig, A. Mufti, *J. Phys. D: Appl. Phys.* **1989**, *22*, 1924.
134. D. Kendall, *J. Vac. Sci. Technol.* **1990**, *A8*, 3598.
135. L. C. Chen, M. Chen, T. H. Tsaur et al., *Sens. Actuators, A* **1995**, *49*, 115.
136. M. Miyashita, T. Tusga, K. Makihara et al., *J. Electrochem. Soc.* **1992**, *139*, 2133.

137. Y. P. Xu, R. S. Huang, *J. Electrochem. Soc.* **1990**, *137*, 948.
138. R. L. Gealer, H. K. Karsten, S. M. Ward, *J. Electrochem. Soc.* **1988**, *135*, 1180.
139. S. A. Campbell, D. J. Schiffrin, P. J. Trfton, *J. Electroanal. Chem.* **1993**, *344*, 211.
140. H. G. Linde, L. W. Austin, *Sens. Actuators, A* **1995**, *49*, 167.
141. Y. Nemirovsky, A. El-Bahar, *Sens. Actuators* **1999**, *75*, 205.
142. M. C. Acero, J. Esteve, Chr. Burrer et al., *Sens. Actuators, A* **1997**, *46*, 22.
143. P. T. Michaud, D. Babic, *J. Electrochem. Soc.* **1998**, *145*(11), 4040.
144. I. Barycka, H. Teterycz, Z. Znamirowski, *J. Electrochem. Soc.* **1979**, *126*, 345.
145. V. Lehmann, K. Mitani, D. Feijoo et al., *J. Electrochem. Soc.* **1991**, *138*, L3.
146. No reference.
147. C. A. Desmond, C. E. Hunt, S. N. Farrens, *J. Electrochem. Soc.* **1994**, *141*, 178.
148. H Seiter, Integrational Etching Methods in *Semiconductor Silicon 1977* (Eds.: H. R. Huff, E. Sirtl), Electrochemical Society, Pennington, NJ, 1977, pp. 187.
149. E. D. Palik, V. M. Bermudez, O. J. Glembocki, *J. Electrochem. Soc.* **1985**, *133*, 135.
150. S. D. Collins, *J. Electrochem. Soc.* **1997**, *144*(6), 2242.
151. G. S. Calabrese, M. S. Wrighton, *J. Electrochem. Soc.* **1981**, *128*, 1014.
152. L. D. Dyer, G. J. Grant, C. M. Tipton et al., *J. Electrochem. Soc.* **1989**, *136*, 3016.
153. Q. Zhang, L. Liu, Z. Li, *Sens. Actuators, A* **1996**, *56*, 251.
154. P. J. Hesketh, C. Ju, S. Gowda et al., *J. Electrochem. Soc.* **1993**, *140*, 1080.
155. D. L. Kendall, *Appl. Phys. Lett.* **1975**, *26*, 195, 287.
156. R. L. Smith, B. Kloeck, S. D. Collins, *J. Electrochem. Soc.* **1988**, *135*, 2001.
157. P. Allongue, V. C. Kieling, H. Gerischer, *J. Electrochem. Soc.* **1993**, *140*, 1009.
158. P. Jakob, Y. J. Chabal, *J. Chem. Phys.* **1991**, *95*(4), 2897.
159. M. Elwenspoek, *J. Electrochem. Soc.* **1993**, *140*, 2075.
160. H. C. Gatos, M. C. Lavine, in *Progress in Semiconductors* (Eds.: A. F. Gibon, R. E. Burgess), CRC Press, Cleveland, Ohio, 1965, pp. 1, Vol. 9.
161. P. Gorostiza, R. Diaz, F. Sanz et al., *J. Electrochem. Soc.* **1997**, *144*, 4119.
162. I. Barycka, I. Zubel, *Sens. Actuators, A* **1995**, *48*, 229.
163. Cs. Ducso, E. Vazsonyi, M. Adam et al., *Sens. Actuators, A* **1997**, *60*, 235.
164. R. A. Lee, C. Patel, H. A. Williams et al., *IEEE Trans. Electron Devices* **1989**, *36*(11), 2703.
165. G. K. Mayer, H. L. Offereins, H. Sandamaier et al., *J. Electrochem. Soc.* **1990**, *137*, 3947.
166. M. J. Declercq, L. Gerzberg, J. D. Meindl, *J. Electrochem. Soc.* **1975**, *122*, 545.
167. R. Divan, N. Moldovan, H. Camon, *Sens. Actuators* **1999**, *74*, 18.
168. W. P. Maszara, *J. Electrochem. Soc.* **1991**, *138*, 341.
169. T. A. Kwa, P. J. French, R. F. Wolffenbuttel et al., *J. Electrochem. Soc.* **1995**, *142*, 1226.
170. K. Ljungberg, U. Jansson, S. Bengtsson et al., *J. Electrochem. Soc.* **1996**, *143*, 1709.
171. R. Takizawa, T. Nakanishi, K. Honda et al., *Jpn. J. Appl. Phys.* **1988**, L2210.
172. G. S. Higashi, Y. J. Chabal, G. W. Trucks et al., *Appl. Phys. Lett.* **1990**, *56*(7), 656.
173. S. Watanabe, M. Shigeno, N. Nakayama et al., *Jpn. J. Appl. Phys.* **1991**, 3575.
174. K. Sawara, T. Yasaka, S. Miyazaki et al., *Jpn. J. Appl. Phys.* **1992**, *31*, L931.
175. Y. J. Chabal, G. S. Higashi, K. Raghavachari et al., *J. Vac. Sci. Technol.* **1988**, *A 7*, 2107.
176. B. Schwartz, H. Robins, *J. Electrochem. Soc.* **1976**, *123*, 1903.
177. Q. B. Vu, D. A. Stricker, P. M. Zavracky, *J. Elechochem. Soc.* **1996**, *143*, 1372.
178. A. Uhlir Jr., *Bell Syst. Tech. J. March*, **1956**, 333.
179. T. Ohmi, M. Miyashita, M. Itano et al., *IEEE Trans. Electron Devices* **1992**, *39*(3), 537.
180. H. Nojiri, M. Uchiyama, *Sens. Actuators* **1992**, *A 34*, 167.
181. C. R. Tellier, A. Brahim-Bounab, *J. Mater. Sci.* **1994**, *29*, 6354.
182. F. Shimura, *J. Electrochem. Soc.* **1980**, *127*, 551.
183. T. Ohmi, T. Imaoka, I. Sugiyama et al., *J. Electrochem. Soc.* **1992**, *139*, 3317.
184. W. Kern, *J. Electrochem. Soc.* **1990**, *137*, 1887.
185. G. J. Norga, M. Platero, K. A. Black et al., *J. Electrochem. Soc.* **1997**, *144*, 2801.
186. H. Shimada, M. Onodera, S. Shimomura et al., *J. Electrochem. Soc.* **1992**, *139*, 1721.

187. D. Tolliver, *Solid State Technol. November*, **1975**, 33.
188. O. J. Anttila, M. V. Tilli, M. Schaekers et al., *J. Electrochem. Soc.* **1992**, *139*, 1180.
189. D. Chopra, I. I. Suni, A. A. Busnaina, *J. Electrochem. Soc.* **1998**, *145*(4), L60.
190. L. Mouche, F. Tardif, J. Derrien, *J. Electrochem. Soc.* **1994**, *141*, 1684.
191. No reference.
192. V. D. Archer, *J. Electrochem. Soc.* **1982**, *129*, 2074.
193. K. H. Yang, in *Semiconductor Silicon 1977* (Eds.: H. R. Huff, E. Sirtl), Electrochemical Society, Pennington, NJ, 1977, pp. 187.
194. W. C. Dash, *J. Appl. Phys.* **1956**, *27*, 1193.
195. P. H. Bellin, W. K. Zwicker, *J. Appl. Phys.* **1971**, *42*, 1216.
196. K. S. Choe, *J. Electrochem. Soc.* **1995**, *142*, 1647.
197. K. R. Mayer, *J. Electrochem. Soc.* **1973**, *120*, 1780.
198. R. W. Series, K. G. Barraclough, W. Bardsley, *J. Cryst. Growth* **1980**, *49*, 363.
199. J. L. Hoyt, E. F. Crabbe, R. F. W. Pease et al., *J. Electrochem. Soc.* **1988**, *135*, 1839.
200. Z. Shi, *J. Electrochem. Soc.* **1995**, *142*, 2377.
201. R. L. Meek, M. C. Huffstutler Jr., *J. Electrochem. Soc.* **1969**, *109*, 893.
202. A. Bohg, *J. Electrochem. Soc.* **1971**, *118*, 401.
203. J. C. Greenwood, *J. Electrochem. Soc.* **1969**, *116*, 1325.
204. C. P. Wu, E. C. Douglas, C. W. Mualler et al., *J. Electrochem. Soc.* **1979**, *126*, 1982.
205. R. Chmolke, G. Puppe, H. Piontek, *J. Electrochem. Soc.* **2000**, *147*, 2999.
206. G. R. Booker, R. Stickler, *J. Electrochem. Soc.* **1962**, *109*, 1167.
207. B. L. Sopori, *J. Electrochem. Soc.* **1984**, *131*, 667.
208. F. Subrahmanyan, H. Z. Massoud, R. B. Fair, *Appl. Phys. Lett.* **1988**, *52*(25), 2145.
209. R. J. Jaccodine, *J. Electrochem. Soc.* **1963**, *110*(6), 524.
210. H. Kasapbasioglu, P. J. Hesketh, *J. Electrochem. Soc.* **1993**, *140*, 2319.
211. B. A. Unvala, D. B. Holt, A. San, *J. Electrochem. Soc.* **1972**, *19*, 318.
212. R. Kinder, *Phys. Stat. Sol.* **1997**, *164*, 785.
213. Y. Matsumoto, T. Shimada, M. Ishida, *Sens. Actuators* **1999**, *A 72*, 153.
214. H. R. Huff, E. Sirtl, (Eds.), Silicon Wafer Process Technology: Slicing, Etching, Polishing in *Semiconductor Silicon 1977*, Electrochemical Society, Pennington, NJ, 1977, pp. 154.
215. W. L. C. M. Heyboer, G. A. C. M. Spierings, J. E. A. M. van den Meerakker, *J. Electrochem. Soc.* **1991**, *138*, 774.
216. M. V. Sullivan, D. L. Klein, R. M. Finne et al., *J. Electrochem. Soc.* **1963**, *10*, 412.
217. J. D. Olivas, S. Bolin, *J. Med. January*, **1998**, 38.
218. O. Than, S. Buttgenbach, *Sens. Actuators, A* **1994**, *45*, 85.
219. P. M. Sarro, A. W. van Herwaarden, *J. Electrochem. Soc.* **1986**, *133*, 1724.
220. C. Levy-Clement, A. Lagoubi, R. Tenne et al., *Electrochim. Acta* **1992**, *37*(5), 877.
221. P. Lim, J. R. Brock, I. Trachtenberg, *Appl. Phys. Lett.* **1992**, *60*(1), 112.
222. A. Gotz, J. Esteve, J. Bausells et al., *Sens. Actuators, A* **1993**, *37*, 744.
223. L. Smith, A. Soderbarg, *J. Electrochem. Soc.* **1993**, *140*, 271.
224. B. Kloeck, S. D. Collins, N. F. De Rooij et al., *IEEE Trans. Electron Devices* **1989**, *36*(4), 663.
225. Y. Zu, L. Xie, B. Mao et al., *Electrochim. Acta* **1998**, *43*(12–13), 1683.
226. V. Lehmann, *J. Electroanal. Chem.* **1993**, *140*, 2836.
227. S. Meltzer, D. Mandler, *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 1019.
228. S. Meltzer, D. Mandler, *J. Chem. Soc., Faraday Trans.* **1995**, *91*(6), 1019.
229. H. K. Kuiken, R. P. Tijburg, *J. Electrochem. Soc.* **1980**, *130*, 1722.

12
Electrochemical Machining

Alexei D. Davydov
Frumkin Institute of Electrochemistry, Russian Academy of Sciences, Moscow, Russia

Vladimir M. Volgin
Tula State University, Tula, Russia

12.1 Introduction **811**
12.1.1 Definition and General Information 811
12.1.2 Advantages and Disadvantages 812

12.2 Electrode Processes **813**
12.2.1 Anodic Processes 813
12.2.2 Cathodic Processes 819

12.3 Metal Removal Rate **820**

12.4 Accuracy **821**
12.4.1 Methods of Lowering the Effects of Void Fraction and Heating on the Distribution of Current Density Over the WP Surface 821
12.4.2 Methods Enhancing the Localization of Metal Dissolution 822

12.5 Prediction of Workpiece Shape and Tool-electrode Design **823**
12.5.1 Problem with Moving Boundary 826
12.5.1.1 The Direct Problem 826
12.5.1.2 The Inverse Problem 827
12.5.2 Quasi-steady-state Approximation 828
12.5.2.1 Calculation of Transfer Processes in the IEG 829
12.5.2.1.1 Calculation of Transfer Processes in the Bulk IEG 829
12.5.2.1.2 Calculation of Transfer Processes in the Near-electrode Diffusion Layer 831
12.5.2.2 Direct Problem: Evolution of WP Surface 831
12.5.2.3 The Inverse Problem: Correction of TE Surface 832
12.5.3 Local, One-dimensional Approximation 833

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

12.5.3.1	Calculation of Transfer Processes in the IEG	835
12.5.3.2	The Direct Problem: The Evolution of WP Surface	837
12.5.3.3	The Inverse Problem: The Correction of TE Surface	837
12.5.4	Automated Design of ECM Operations	838
12.6	**ECM Techniques and Applications**	**839**
12.6.1	Electrochemical Sinking	839
12.6.2	Electrochemical Drilling	841
12.6.3	Electrochemical Deburring	845
12.6.4	Removing of Defective Layers	845
12.6.5	ECM in Biomedical Engineering	846
12.6.6	Subsea Application	847
12.6.7	Electrochemical Micromachining (ECMM)	847
12.6.8	Hybrid Processes	850
12.7	**Surface Layer Integrity after ECM**	**850**
12.8	**Control of Electrolyte Composition and Environmental Problems**	**851**
12.9	**Conclusions**	**852**
	References	852

12.1 Introduction

12.1.1 Definition and General Information

Electrochemical machining (ECM) is a method of metal machining that aims at producing parts of specified shape, dimensions, and surface finish. The process is based on the removal of metal by electrochemical dissolution ([1–15] and references cited therein). Special machines have been developed to realize this aim. A complete ECM installation (Fig. 1) consists of the machine, the power supply, the electrolyte circulation system (tank, pump, heat exchanger, and sludge removal unit), and the control system (control of current, voltage, feed rate, gap width, and electrolyte temperature, pH value, pressure, and concentration; short-circuit protection).

The process is conducted in the working chamber (electrochemical cell) of the machine, where a workpiece (WP) and a tool electrode (TE) are placed. The WP is connected to the positive pole of a power supply, and the TE serves as the cathode. The interelectrode distance is typically 0.02–0.8 mm. The electrolyte (usually an aqueous solution of an inorganic salt, 15% $NaNO_3$ or NaCl, for example) is pumped through the interelectrode gap (IEG) at a velocity of 5–50 m s^{-1}, in order to remove the electrode reaction products (gases, hydroxides) and the heat generated by the current. Current densities of 10 to 150 A cm^{-2} provide for the high productivity of the process. The voltage is typically 6–18 V. In most cases, the TE is moved in the direction of the intended anodic dissolution (metal loss) in order to maintain a prescribed machining gap. As the anode dissolves, the TE shape is reproduced on the WP (Fig. 2). A smaller gap width provides more accurate reproduction and a higher dissolution rate.

During ECM, electrochemical dissolution of anode and cathodic evolution of hydrogen proceeds on the electrodes (the WP and TE, respectively). Along with these basic reactions, parallel reactions proceed concurrently, for example, oxygen anodic evolution, cathodic reduction of nitrate ions, if $NaNO_3$ electrolyte is used. It is important to note that electrochemical reactions in a narrow IEG result in gas evolution. The temperature of the electrolyte in the IEG and the void fraction increase as the electrolyte flows along the gap. This leads to a variation in the electrolyte conductivity that has an effect on the distributions of current and metal dissolution rate over the WP surface. The electrode processes and the processes in

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

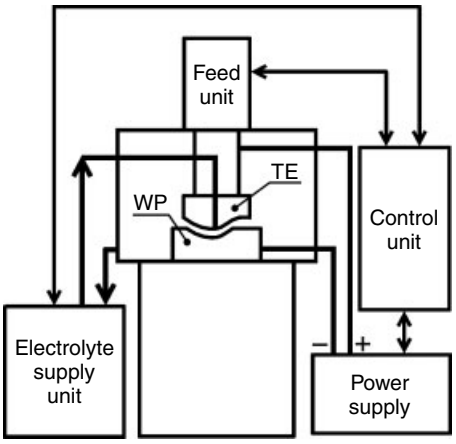


Fig. 1 Schematic of ECM installation.

the IEG are considered in the following sections.

The ECM finds use in the aerospace and other industries for manufacturing turbine blades of engines and dies, deep small-diameter hole drilling, deburring, and other processes.

The development of ECM started in the early twentieth century. The Russian engineer V.N.Gusev was the first to take out a patent for all basic principles of ECM [16]. However, Gusev used rather high voltages (of an order of 100 V). Therefore, it is possible that, in his work, along with the electrochemical metal removal, electrodischarge removal also occurred. Before Gusev's patent, in the very early part of the twentieth century, a study [17] on the electrolytic drilling and slotting was published, which escaped notice and, it seems, has not been cited until now. Cowper-Coles [17] showed that holes of any shape can be drilled in hardened armor plates by anodic dissolution in circulating H_2SO_4 or NaCl solutions. Some information on the history of the development of industrial ECM in the middle of the twentieth century in the United States and Great Britain is available in Refs 10 and 11.

12.1.2 Advantages and Disadvantages

The absence of any physical contact between the TE and the WP is the reason for both the advantages and disadvantages of the method. The ECM productivity, as well as the productivity of any other electrochemical process, is governed by Faraday's and Maxwell's laws and is independent of the mechanical properties of the WP material, in particular, hardness. Therefore, ECM is feasible for the machining of difficult-to-cut materials (hardened steels, superalloys, magnetic alloys, Ti, W, WC, and so forth). The absence of cutting forces virtually eliminates the tool wear and allows one to machine fragile materials and thin-walled parts. The TE may be of various designs that enable machining in the difficult-to-reach areas, inside a complex-shaped part, for example. After ECM, the product is free from burrs. Three-dimensional products can be machined in a single step. With the application of ECM, it is possible to generate more freedom of design for the product. In contrast to other methods of machining, in ECM, as a rule, decreasing roughness of the surface is

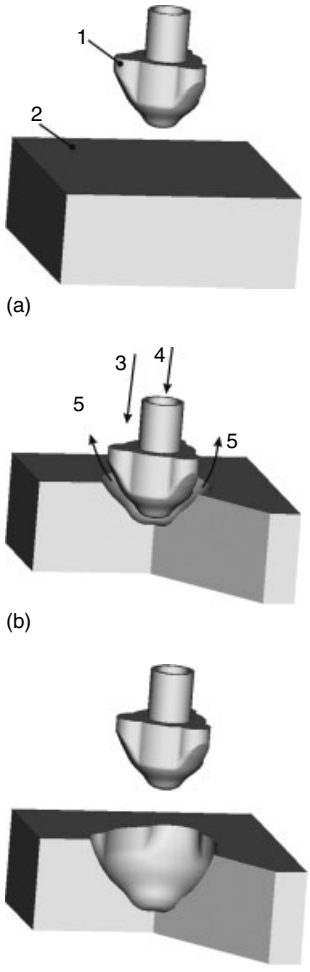


Fig. 2 Schematic illustration of ECM: TE and WP (a) before the ECM, (b) in the course of ECM, and (c) after the ECM: (1) the TE, (2) the WP, (3) the feed rate, (4) the electrolyte inlet, and (5) the electrolyte outlet.

observed with a simultaneous increase in productivity.

On the other hand, the existence of a gap between the TE and the WP reduces the accuracy of TE reproduction. The distribution of metal dissolution rate over the machining surface is determined by the distribution of the electric field in the complex-shaped interelectrode space, which is filled with an electrolyte with varying conductivity. The electrolytes cause corrosion of equipment. In some cases, the TE is complex and expensive. A special TE has to be developed for each product; therefore, the ECM processes are best suited to large-lot production applications.

The ECM finds use in the production of gas turbine engine parts and for other aerospace applications, in general manufacturing applications [1, 3, 8, 10, 11, 13, 14, 18, 19, 20]. Some of them are considered in more detail in Sect. 12.6.

12.2 Electrode Processes

12.2.1 Anodic Processes

High-rate anodic dissolution of metals has been studied intensively (Reviews [9, 12, 14, 21]). Several types of anodic metal dissolution are recognized. Sometimes, the dissolution of metals in the active state is used in ECM. Deep, small-diameter holes drilled in acidic solutions are an example. An acid is used in order to avoid sludge formation, which hampers the drilling. However, highly aggressive acidic solutions cause the corrosion of equipment and are hazardous to the maintenance personnel. In most cases, aqueous salt solutions with so-called activating anions are applied for ECM. The Cl^- and Br^- anions are the strongest activators and NO_3^- is a weaker one. Activating electrolytes are used in order to overcome the metal tendency to passivate and hence to achieve high rates of anodic metal dissolution. After a shift of the metal potential, $\Delta E_a = E^{\#} - E_{oc}$, in the

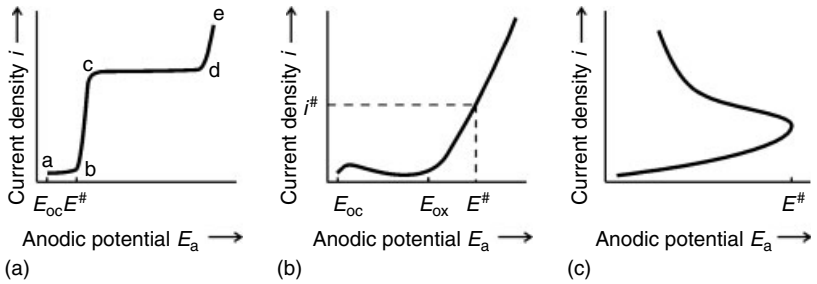


Fig. 3 Anodic voltammograms typical of ECM (for comments see the text).

anodic direction from the open-circuit potential, E_{oc}, to the activation potential, $E^{\#}$, passivity breaks down and the current increases steeply (the voltammogram on Fig. 3a). If the breakdown of only the initial passivity is required, ΔE_a may comprise only tens of millivolts (Al, Fe in NaCl solutions). In other cases, prior to the onset of activation, the anodic oxide film grows, and ΔE_a may comprise a fraction of a volt or even several volts (Al or Be in NH_4NO_3, Nb in KBr, Ti alloys in NaCl). The processes of passivity breakdown and activation propagation from the initial centers to the entire WP surface as applied to ECM were considered in Ref. 9.

The rate of active metal dissolution is restricted by the limiting current. Two main types of the limiting current of high-rate anodic metal dissolution are recognized. The limiting current of the first type is reached when the near-electrode concentration of the reagent (generally electrolyte anions) decays to zero (tungsten dissolution in alkali, for example). The limiting current of the second type is reached when the concentration of the dissolution product (commonly, metal salt) increases to saturation (the dissolution of Ni or Fe in NaCl). When the dissolution proceeds in the limiting-current range (cd in Fig. 3a), the metal is covered with a film of the products of anodic reaction. The limiting current depends on the electrolyte flow rate in the gap and the solution concentration at the inlet to the gap; therefore, an increase of the flow rate and the optimization of concentration are the principal ways of raising the limiting current. A review, Ref. 22, is devoted to an analysis of the problem of the limiting currents of metal anodic dissolution. The potential range of the limiting current cd (Fig. 3a) depends on the properties of the anodic film and the parameters of the process (the flow rate, the concentration, and the electrolyte temperature). When a certain value of the anodic potential is exceeded, the current may increase again: the overlimiting current range de is reached.

The case in which the electrolyte anions have a relatively weak activating action, for example NO_3^- on nickel, is important for ECM. A special feature of this case is that $E^{\#}$ is significantly higher than in NaCl and, most importantly, higher than the potential at the onset of anodic oxygen evolution E_{ox} (Fig. 3b). The use of weakly activating electrolytes enables one to enhance the localization of metal dissolution on the areas with the smallest IEG and the highest current density i ($i > i^{\#}$, Fig. 3b) and to reduce the metal dissolution on the WP areas adjacent to the area to be machined (overcutting). On the adjacent areas, the current flows, but it is lower than $i^{\#}$ (the gap is larger) and, hence, is consumed by the

oxygen evolution rather than by metal dissolution. In this case, a high degree of localization of metal dissolution results from an abrupt decrease in the current efficiency at $i < i^{\#}$. The position of the boundary on the WP surface between the area where the machining is required and the adjacent area (between the areas with $i > i^{\#}$ and $i < i^{\#}$) depends on the ECM parameters (voltage, temperature, electrolyte concentration, and flow rate). For instance, to achieve a higher degree of localization, the voltage should not exceed the minimum value that is necessary to provide the required productivity.

The transition to the pulse conditions with proper ECM parameters enables one to raise $i^{\#}$ and, thus, further enhance the degree of localization of metal dissolution [9, 23]. At the optimal (sufficiently short) pulse-on time, the activation process does not have a period that is sufficiently long for high-rate metal dissolution to start on the adjacent areas of the WP surface, even if $E_a > E^{\#}$ in these areas. This is due to the fact that intense metal dissolution does not begin immediately upon imposing a voltage, but rather does so only after a certain charge Q^* has passed, which is consumed for the preliminary processes: charging of the electrical double layer, the formation of an oxide film with an increase in the potential from E_{oc} to $E^{\#}$, and oxygen evolution in the range $E_{ox} < E_a < E^{\#}$. Under optimal pulse conditions of ECM, only in the working zone (in the area with the minimum IEG), the charge in the pulse exceeds Q^* and leads to intense metal dissolution. In the adjacent areas, where IEG is larger, the entire charge in a pulse will be consumed only for the aforementioned preliminary processes. High-rate dissolution is concentrated only in the area of the smallest IEG. Figure 4 schematically shows the effect of electrolyte composition and the ECM conditions on the shaping process.

To analyze the dissolution of some metals, for example titanium, it is advantageous to examine galvanostatic (galvanodynamic) curves rather than potentiostatic (potentiodynamic) data. This is associated with the fact that the anodic metal activation occurs at high $E^{\#}$ when an anodic oxide film having a low electron conductivity is rather thick. The breakdown of this highly resistive film by activating anions at $E^{\#}$ leads to a decrease in E_a (Fig. 3c). The ECM proceeds at $E_a < E^{\#}$, and the galvanostatic experiments enable one to determine this value of E_a.

The composition of the alloy to be machined is important in the analysis of the problem of machinability by ECM. A difference in the tendencies of different alloy components to passivate in the electrolyte of a given composition is of prime importance. If the content of component A, which is less prone to passivate, is not too high, it is etched away from a thin alloy surface layer and the potential increases to a value at which high-rate dissolution of the more passivation-prone component B is possible. Further alloy dissolution proceeds at this potential. The surface is enriched with component B, and both components dissolve concurrently.

If the content of component A is very high, two cases can occur. (1) Component B dissolves simultaneously with A at relatively low potentials. This is observed in the high-rate dissolution of two-component nickel–chromium alloys. The anodic behavior of Ni−Cr alloys will be considered in more detail in the following paragraphs. (2) Only component A dissolves electrochemically at the rather low potential required for its dissolution. Particles of component B crumble from the anode

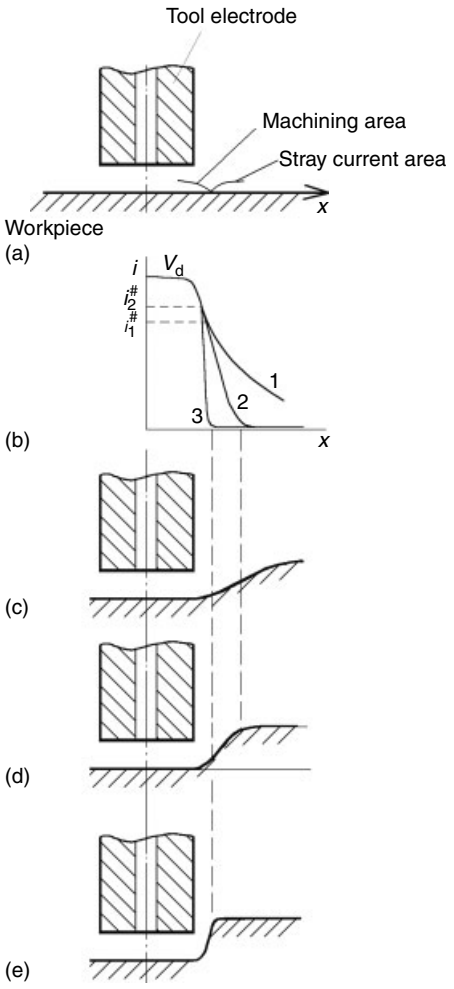


Fig. 4 Schemes of ECM explaining different degrees of anodic dissolution localization: (a) the ECM scheme, (b) the distributions of the current density, i, and the dissolution rate, V_d, over the machining area and the adjacent area (stray-current area) under various conditions of ECM, (c,d) shapes of cavities produced by the ECM with the same tube TE under various conditions of ECM. (1) The distributions of current density i during the ECM in NaCl and $NaNO_3$ solutions and the dissolution rate V_d in the NaCl solution. (2) The distributions of the dissolution rate, V_d, in ECM with a constant voltage in the $NaNO_3$ solution. (3) The distribution of dissolution rate during pulse ECM in $NaNO_3$ solution.

surface. An example is the high-rate anodic dissolution of multicomponent superalloys based on Ni and Cr, where the γ phase crumbles from the surface [24].

However, sometimes component B is passive over too wide a potential range and its content in the alloy is high; hence ECM becomes practically impossible. In this case, the electrolyte should be replaced by one in which the range of B passivity is not so wide, or a solution composition should be chosen involving, for example, two components that provide for high-rate dissolution of both components. A mixture of NaOH and $NaNO_3$ is an example of the use of two-component electrolyte for the machining of two-component tungsten carbide–cobalt alloy [25]. At the optimal solution composition, the alkali provides the high-rate dissolution of WC and does not hinder Co dissolution (cobalt is passive in alkali) and the nitrate provides the high-rate dissolution of cobalt binder but does not passivate WC. In doing so, minimum roughness along with a high productivity of ECM is attained.

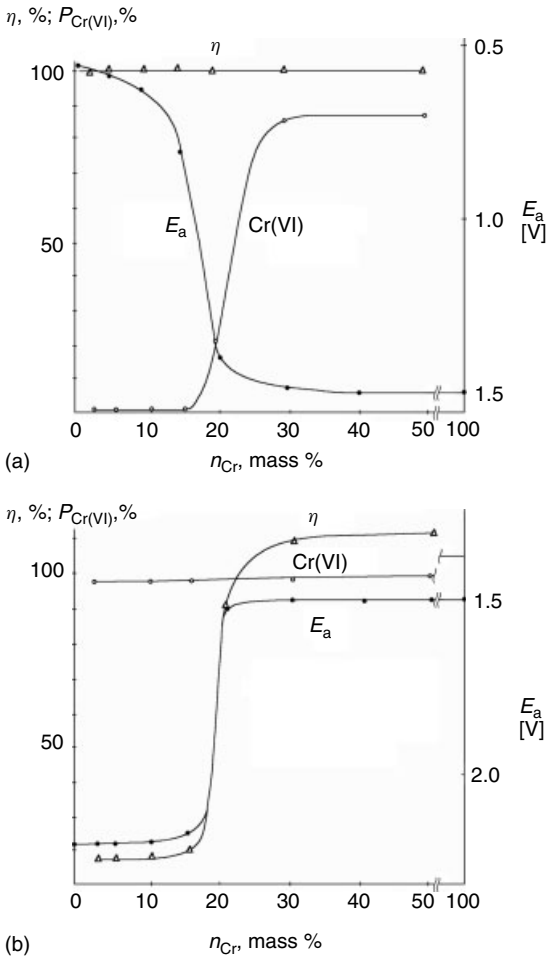


Fig. 5 Plots of the anodic potential, E_a, at which Cr, Ni, and their alloys dissolve at a current density of 8 A cm^{-2}, the current efficiency η, and a Cr(VI) fraction of the total Cr amount in the solution versus the chromium content in the alloy n_{Cr}. The experiments were performed in (a) 2 M NaCl solution and (b) 1 M Na_2SO_4 (the results obtained in Na_2SO_4 and $NaNO_3$ solutions are similar) on the rotating (500 rpm) disk electrodes.

The regularities of high-rate anodic dissolution of many materials were considered in Ref. 9.

As an example, the experimental results for the anodic dissolution of alloys of the nickel–chromium system are presented in Figs 5 and 6. This composition is the basis of several superalloys, which are machined using the ECM. Alloy components – chromium and nickel – exhibit different tendencies to passivate. Nickel is weakly passivated in the NaCl solution; its tendency to passivate is much stronger in the $NaNO_3$ and Na_2SO_4 solutions. Chromium is passive in all of these solutions at potentials up to 1.5 V (SHE). In the NaCl solutions, the activation and intense dissolution of nickel proceed at less positive potentials than the potential of chromium passivity breakdown. In the $NaNO_3$ and Na_2SO_4 solutions, the anodic dissolution on nickel proceeds at significantly more positive potentials than the potential of chromium dissolution; therewith, in these solutions (especially in Na_2SO_4), a large fraction of the current is

consumed not for metal dissolution but for the parallel process of oxygen evolution.

With an increase in the chromium content of the alloy, the alloy's behavior changes from the behavior of pure nickel to that of pure chromium (Fig. 5). It is important that this change proceeds jumpwise in a rather narrow range of alloy composition between 15 and 20 wt% of chromium [26].

Chemical analysis of the solutions after anodic dissolution have shown that the oxidation state of chromium in the dissolution products depends on the alloy composition and, correspondingly, on the potential of alloy dissolution. At potentials less positive than the potential of the onset of pure-chromium passivity breakdown, chromium dissolves from the nickel-based alloys as Cr(III). The alloys with chromium contents of not more than 15% dissolve in this manner in NaCl solution. At higher E_a, chromium from the alloy dissolves, for the most part (about 90%), in the form of Cr(VI). This is true for all alloys in Na_2SO_4 (or $NaNO_3$) solution and for the alloys containing more than 25% chromium in NaCl solution.

The current efficiency of alloy dissolution is calculated by the following equation:

$$\eta = \frac{\Delta m F}{Q} \Sigma \frac{n_i z_i}{A_i} \qquad (1)$$

where Δm is the alloy mass that dissolves when a charge Q is passed through the electrode, n_i is the mass fraction, A_i is the atomic mass, and z_i is the oxidation state of the ith component, respectively. In the case of the anodic dissolution of nickel–chromium alloys, the equation for η should be written in the following form:

$$\eta = \frac{\Delta m F}{Q} \left[\frac{n_1 z_1}{A_1} + \frac{n_2 q_1 z_2}{A_2} + \frac{n_2 q_2 z_3}{A_2} \right] \qquad (2)$$

where A_1 and A_2 are the atomic masses of Ni and Cr, respectively; n_1 and n_2 are the mass fractions of Ni and Cr in the alloy, respectively; $z_1 = 2$, $z_2 = 6$, and $z_3 = 3$ are the oxidation states of Ni (2) and Cr (6 and 3) in the compounds that form in the alloy anodic dissolution process; and q_1 and q_2 are the fractions of Cr(VI) and Cr(III) in the solution, respectively, which are determined by chemical analysis ($q_1 + q_2 = 1$). To obtain the value of η in terms of percent, the right-hand side of Eq. (2) should be multiplied by 100.

In NaCl solutions, the current efficiency is 100% for the dissolution of all Ni−Cr alloys, as well as for pure nickel and pure chromium. In sulfate and nitrate electrolytes, η is low for alloys that are nickel rich, because nickel itself dissolves with a low current efficiency under these conditions (the main fraction of the current is consumed by oxygen evolution).

Figure 6 displays plots of alloy dissolution productivity (P) versus chromium content for two different electrolytes at a current density of 8 A cm^{-2}. The maximum in Curve 2 is associated with an effect of chromium content of the alloy on the current efficiency in Na_2SO_4 and $NaNO_3$ solutions. The variations in the electrochemical equivalent of the alloy and in the current efficiency with an increase in the Cr content have the strongest effect on the dissolution productivity. An increase in the oxidation state of chromium at $n_{Cr} \approx 15\%$ has a weak effect on the productivity (a small inflection in the Curve 1 for NaCl). The shape of Curve 2 in Fig. 6 for the $NaNO_3$ solution depends on the current density: in the range of low current densities, the productivity (P) of ECM of nickel and nickel-rich alloys increases with the current density, owing to an increase in the current efficiency.

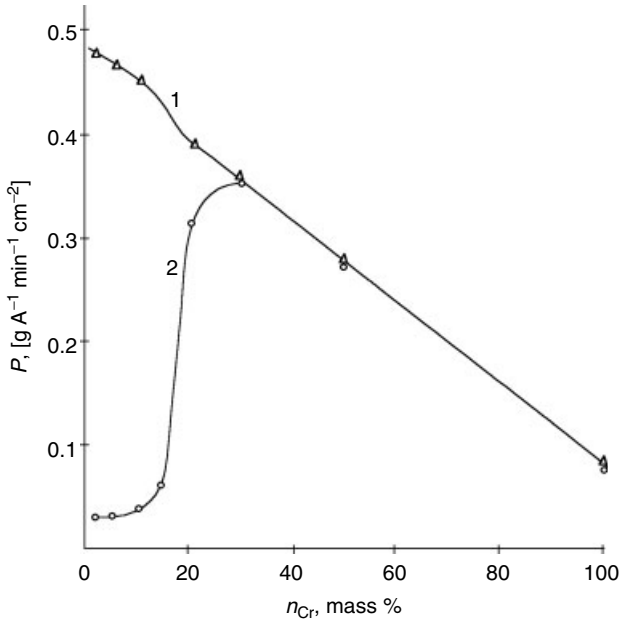


Fig. 6 Plots of anodic dissolution rate, P, in weight units, of nickel, chromium, and their alloys in (1) NaCl and (2) Na_2SO_4 solutions (the results obtained in Na_2SO_4 and $NaNO_3$ solutions are similar) at 8 A cm^{-2} versus the chromium content in the alloy n_{Cr}. (The plots are obtained using the data presented in Fig. 5).

12.2.2 Cathodic Processes

In NaCl solutions, one of the typical ECM electrolytes, the electrochemical decomposition of water with the formation of hydrogen gas is the only cathodic reaction:

$$2H_2O + 2e \longrightarrow H_2\uparrow + 2OH^- \qquad (3)$$

The kinetics of this reaction on stainless steel cathode in a wide range of current densities up to very high i can be described by the Tafel equation $E_c = -1.25 - 0.14 \log i$ [27], where E_c (in volts) is referred to a standard hydrogen electrode (SHE), i is expressed in amperes per square centimeter.

An important consequence of this cathodic reaction is the fact that the IEG becomes filled with a two-phase, gas–liquid mixture. This, along with the electrolyte heating with the flowing current, leads to the variation of the electrolyte conductivity in the gap and, correspondingly, of the current density at the anode. The latter makes the production of a part with prescribed sizes by the ECM method difficult. In the case of too large a void fraction (it depends on the current density, the electrolyte flow rate, the pressure, and so forth), local electrical breakdown within the gap (sparking) may occur, which can degrade the surface finish and damage the TE surface [28].

In the nitrate electrolyte, the reduction of nitrate ions takes place on the cathode. The complete reduction of nitrate in the neutral

and alkaline solutions yields ammonia:

$$NO_3^- + 7H_2O + 8e \longrightarrow NH_4OH + 9OH^- \quad (4)$$

However, in specific cases, the degree of reduction depends on the electrochemical conditions. For example, the first reduction stage

$$NO_3^- + H_2O + 2e \longrightarrow NO_2^- + 2OH^- \quad (5)$$

occurs at potentials that are not as negative as those of hydrogen evolution. The rate of nitrate reduction is restricted by the limiting current that is low compared to the high current commonly applied in ECM. After the limiting current is reached, the reactions of nitrate reduction and hydrogen evolution proceed simultaneously, obeying the kinetic equation $E_c = -1.19 - 0.14 \log i$ [27].

More detailed information on the cathodic processes in ECM is available from a book [9] and references cited therein.

12.3 Metal Removal Rate

The rate V_a of recession of the WP surface is determined by the equation

$$V_a = \frac{\eta\varepsilon}{\rho} i \quad (6)$$

where η is the current efficiency; ε is the electrochemical equivalent; and ρ is the metal density. In some cases, it is difficult to calculate ε, because the number of electrons in the dissolution reaction is uncertain. This is true for metals such as Al, Mg, Be, Fe, Cr, Cu, and some others. The number of electrons for the same metal may depend on the electrolyte composition, the anodic potential, and, for alloys, also on the metal content in the alloy. Frequently, the determination of ε requires the experimental data. The current efficiency η for the dissolution reaction may depend on E_a (or i). In the simplest case, the current density is determined by the equation

$$i = \frac{\chi(U - E_a + E_c)}{S} \quad (7)$$

where S is the interelectrode gap and U is the voltage between the electrodes. The specific conductivity χ of the electrolyte in the gap depends not only on the electrolyte concentration and temperature at the inlet to the gap but also on the extent of gas evolution and heating in the gap. The electrode potentials are determined experimentally. For example, for the nickel dissolution in nitrate electrolyte at 100 A cm^{-2}, $(E_a + E_c) \cong 3.85$ V.

Metal passivity may significantly raise the E_a and hamper the realization of the high current density that is required for ECM. One way of overcoming passivity, namely the use of activating electrolytes, was considered above. Another popular method of depassivation is the mechanical action on the WP surface in order to remove passive films. This is commonly realized using a moving (rotating, for example) TE with abrasive grains. This method enables one to significantly raise the rate of the ECM of materials that are difficult to activate.

For this purpose, bipolar current can also be used.

The metal dissolution rate in the ECM varies over a wide range, but frequently it is within 0.1–1.0 mm min^{-1}. The impediments to the ECM productivity are frequently associated not with the electrode reactions, but with the processes in the interelectrode space. In the design of the ECM processes, especially for

the machining of large, sculptured surface areas, the problem of the proper arrangement of inlet holes (normally in the TE) to inject the electrolyte into the interelectrode space is of prime importance [3, 10]. The problems of optimization of voltage, electrolyte flow rate, and the TE feed rate are also of importance. Theoretical and experimental studies, and ECM practice, show that the process should be conducted at the minimum required U and with the highest possible feed rate. This allows one to maintain a very small IEG, to obtain high productivity, machining accuracy, and good surface finish. However, there are some impediments to the enhancement of feed rate, primarily electrolyte overheating, leading to boiling. Very intense heating and hydrogen generation in the gap can cause the onset of unwanted electrical discharges, leading to poor surface finish on the WP and the electroerosive wear of the TE. Calculation of the temperature is complicated by the fact that the temperature varies along the flow path as well as across the gap. At present, a 2D model of the temperature distribution has been developed, on the basis of which computer simulations have been generated in order to determine the critical conditions leading to discharges [29]. Using this model, the relationship has been determined between the critical local temperature in the gap and the average temperature at the outlet, which is easily measured and used as a signal for a control system for ECM.

12.4 Accuracy

Two lines of development of methods for providing a highly accurate ECM are as follows:

1. The development of methods for calculating the distribution of local rates of metal dissolution over the machining surface and, consequently, the distribution of gap width. The methods of ECM modeling, which enable one to determine the WP shape after the machining with known TE or, conversely, to determine the TE shape and the machining parameters for the production of the required WP, are discussed in Sect. 12.5.
2. The elaboration of ECM conditions, under which the TE shape is sufficiently accurately reproduced on the WP. Two types of methods are used for this purpose (see Sects 12.4.1 and 12.4.2).

12.4.1 Methods of Lowering the Effects of Void Fraction and Heating on the Distribution of Current Density Over the WP Surface

Among these methods, pulse and pulse-cyclic methods of ECM are of great importance. The voltage pulse duration is chosen such that the conductivity of medium in the gap has no time to change substantially, and the pulse-off time assures the complete renewal of the electrolyte in the gap. Pulse ECM enables one to enhance the uniformity of current density distribution, but reduces the ECM productivity, because a considerable time is taken by the pauses. A decline in the productivity may become intolerably great when a large surface area is being machined: the larger the distance passed by the electrolyte along the gap, the shorter should be the pulse-on time and the longer should be the pulse-off time. To reduce this drawback, the pulse-cyclic conditions of ECM have been elaborated. In this case, the voltage pulse is applied when the gap is small and, in the pause, the TE is removed

from the WP for a larger distance for fast ''washing'' of interelectrode space. By this means, the pulse-off time is reduced. A rich variety of pulse-cyclic conditions of ECM has been developed. The machining with a vibrating electrode has become particularly popular. To reduce the pulse-off time, bipolar pulse modes are also applied [30, 31].

There are some other methods of lowering the effects of void fraction and heating on the distribution of current density over the WP surface.

The pressure P_{out}, which is produced at the outlet, enables the current distribution over the WP surface to be changed. This is associated, to a large degree, with the fact that the gas bubble size decreases with an increase in P_{out}. The void fraction and its effect on the current distribution decrease. At a certain ratio between the inlet pressure and P_{out}, the nonuniformity of the current distribution may be significantly reduced.

The application of air–electrolyte mixtures as a working medium for ECM enables one to raise the localization of metal dissolution in places with the smallest gaps and, thus, to enhance the accuracy of electrochemical reproduction of TE on the WP. To achieve the highest efficiency of this method, several conditions should be fulfilled. The air–electrolyte mixture should be formed in the immediate vicinity of the gap inlet and the flow rate should be adequately high. A ratio between gas and liquid amounts from 3.0 to 3.5 is considered to be most preferable. Rigid stabilization of all the process parameters is required. The design and size of all parts of the mixing device and the values of inlet and outlet pressures are important. To avoid a considerable decline in productivity, the main metal stock should be removed in the air-free or poor-in-air electrolyte. The final stage should be conducted with the highest possible air content. In many cases, the application of air–electrolyte mixture improves the surface finish. For example, surface defects of the striation type can be eliminated.

Sectional TEs involve a set of sections that are separated by insulating layers. By a special control device, the cathodic sections are connected up to the power source alternately in a certain order. By doing so, one can attain a uniform distribution of metal removal rate over the WP surface. Moreover, a comparatively low-amperage power source may be used, and ECM may be performed with a smaller gap. However, owing to the alternate connection of sections, the ECM productivity is low. Another drawback is the etching of nonoperating cathodic sections adjacent to the operating one because of the bipolar effect. In the design of the TE and the engineering process, the following problems should be solved. How many sections should be made and what should be their size? How long does each section work? What is the order of section connection? What is the thickness of the intersection insulator?

12.4.2 Methods Enhancing the Localization of Metal Dissolution

In this group, the method of control of current efficiency distribution is of prime importance (see Sect. 12.2.1), but not just the current density, as in the above group of methods. The mechanism of enhancing a degree of localization of metal dissolution via a proper choice of electrolyte composition and conditions of ECM was considered in detail in Refs 9 and 23.

Among other methods of raising the degree of metal dissolution localization, the electrochemical protection of surface areas immediately adjacent to the region where the dissolution is required should be noted.

Various methods used to improve ECM dimensional control were considered in more detail in Refs 23 and 32.

12.5 Prediction of Workpiece Shape and Tool-electrode Design

In ECM, the WP shape and dimensions vary under the action of the field of dissolution rates, $\boldsymbol{V}_a$. It is normally assumed that the WP anodic dissolution proceeds in the direction of inward normal to its surface, that is, $\boldsymbol{V}_a = \boldsymbol{n}^a V_a$, where $\boldsymbol{n}^a$ is the unit vector of inward normal to the WP surface and V_a is the modulus of anodic dissolution rate (Fig. 7).

In the general form, the physicochemical processes in the IEG are mathematically described as follows:

$$\boldsymbol{V}_a = \frac{\eta\varepsilon}{\rho}\boldsymbol{i} \text{ on } A^t$$

$$\boldsymbol{i} = F\sum_{k=1}^{N_a} Z_k \boldsymbol{J}_k(\boldsymbol{X}) \text{ on } A^t$$

$$\frac{\partial \boldsymbol{X}}{\partial t} = -\text{div}(\boldsymbol{J}_X) + \boldsymbol{q}_X \text{ in } \Omega^t \qquad (8)$$

where F is the Faraday constant; z_k and $\boldsymbol{J}_k$ are the charge and the mass flux of the kth component of the WP material through the anode–electrolyte interface, respectively; $\boldsymbol{X}$ is the vector of generalized spatial fields (phase densities, momentum, hydrodynamic velocity, pressure, temperature, concentrations of electrolyte components, and so forth); $\boldsymbol{J}_x$ is the flux of vector $\boldsymbol{X}$ elements, which is caused by convection, diffusion, and migration; $\boldsymbol{q}_x$ is the intensity of bulk source of vector $\boldsymbol{X}$ elements, for example, the bulk heat generation or a chemical reaction.

The first equation of system (8) expresses Faraday's law; the second equation defines the current density at the WP surface in terms of a sum of mass fluxes of dissolved material cations through the anode/electrolyte interface. The interface mass fluxes at the WP surface depend on the distributions of the electrical field, the concentrations of electrolyte components,

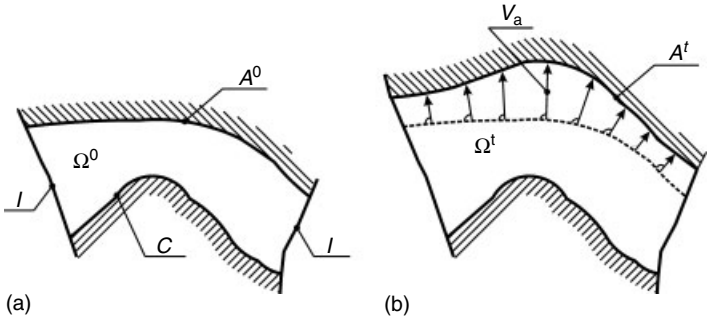


Fig. 7 Scheme of ECM: A^0 is the initial WP surface; A^t is the WP surface during the course of machining; C is the TE surface; Ω^0 and Ω^t are the interelectrode space bounded by the surfaces of the electrodes (C and A) and insulators (I) before the machining (a) and during the course of machining (b).

the temperature, the hydrodynamic velocity, and so forth, which are denoted by the vector of generalized spatial fields $\boldsymbol{X}$, over the IEG. These fields vary in the region Ω^t with moving boundary (Fig. 7) according to the third equation of system (8), which represents the conservation laws for the components of vector $\boldsymbol{X}$, that is, the conservation laws of mass, momentum, energy, and so forth. Depending on the method of describing the WP surface, the evolution of the WP surface is determined by the equations presented in Table 1.

The mathematical description of ECM can be used for solving the *direct* and *inverse* electrochemical shaping problems [4, 9]. In the first case, there is a need to determine the shape and dimensions of the machined WP surface. In this case, the geometry of the TE working surface is prescribed. In the second case, there is a need to determine the shape and dimensions of the TE working surface, in order to produce the WP surface with a required geometry.

In the general case, in order to solve the direct and inverse ECM problems, it is necessary to obtain the simultaneous solution of the basic equations describing the WP surface evolution [9] and a set of equations for the electrode reactions kinetics and transfer processes in the gap, which determines the distribution of current density over the WP surface. Since the shape of one of the electrodes is varied during the machining, as is not known in advance, these problems belong to the class of problems with moving (free) boundaries, and their solution involves great difficulties [4–9]. Therefore, approximate quasi-steady-state and local, one-dimensional methods, which enable one to reduce the ECM problems to those of known boundaries, are widely used [1–9].

In the quasi-steady-state approximation, which is also known as the *step method* [9], it is assumed that the rate of variation in the WP shape, that is, the anodic dissolution rate, is small compared with the rates of transfer processes in the gap; therefore, for calculating the distribution of the current density, the WP surface can be considered as being immobile. This approximation can be used at not very high current densities. At very high current densities, ignoring the WP surface motion during anodic dissolution and the hydrodynamic flow induced by this motion causes a considerable error in the calculated distribution of current density [33].

In the local, one-dimensional approximation [9], the determination of the shape of one electrode is reduced to the determination of the width of the IEG in the machining zone. For instance, if the TE surface is given by the vector function

Tab. 1 Mathematical models of the WP surface evolution at ECM

Method of description	*WP surface*	*Equation of WP surface evolution*
Implicit coordinate	$\Phi(x, y, z, t) = 0$	$\partial\Phi/\partial t + V_a\lvert\text{grad }\Phi\rvert = 0$
	$T(x, y, z) - t = 0$	$V_a\lvert\text{grad }T\rvert = 1$
Explicit coordinate	$z = z(x, y, t)$	$\partial z/\partial t + V_a\sqrt{1 + \lvert\text{grad }z\rvert^2} = 0$
Parametrical	$\Gamma = \Gamma(u, v, t)$, $\Gamma = (x, y, z)$	$\partial\Gamma/\partial t = \boldsymbol{n}^a\, V_a$

$\boldsymbol{r}^{\mathrm{c}}(u, v)$, the WP surface $\boldsymbol{r}^{\mathrm{a}}(u, v, t)$ can be determined as follows:

$$\boldsymbol{r}^{\mathrm{a}}(u, v, t) = \boldsymbol{r}^{\mathrm{c}}(u, v) + S(u, v, t) \cdot \boldsymbol{n}^{\mathrm{c}}(u, v) \tag{9}$$

where u, v are the TE surface parameters; t is the time; $S(u, v, t)$ is the gap width; and $\boldsymbol{n}^{\mathrm{c}}$ is the unit vector of the outward normal to the TE.

In the case of the *inverse problem*, the following equation can be used to determine the TE surface:

$$\boldsymbol{r}^{\mathrm{c}}(u, v, t) = \boldsymbol{r}^{\mathrm{a}}(u, v) + S(u, v, t)\boldsymbol{n}^{\mathrm{a}}(u, v) \tag{10}$$

where $\mathbf{n}^{\mathrm{a}}$ is the unit vector of the outward normal to the WP surface.

The local, one-dimensional approximation may be used if the TE and WP surfaces differ only slightly and a one-to-one correspondence between them can be found. These conditions are fulfilled, if the gap is significantly (by three to five times) wider than the minimum radius of electrode surface curvature. In this case, the electric field in the gap is quasi-uniform; this enables one to express, in an explicit form, the current density at the WP surface in terms of the gap width.

However, even the use of these approximations does not enable one (except for some specific cases) to obtain analytical solutions of the ECM problem. This is due to the significant variations in the properties of the solution in the gap, which are caused mainly by intense heat and gas evolution.

When solving the direct and inverse ECM problems, it is necessary to take into account the peculiarities of the shaping scheme used and the machining conditions. In most cases of ECM, the WP surface is formed under steady-state or nonsteady-state conditions according to one of the following schemes.

1. ECM is performed simultaneously over the entire WP surface in one pass of the contoured TE generally along a single coordinate under the given conditions (voltage, TE feed rate, etc.). The shape and dimensions of TE are reproduced on the WP via its anodic dissolution within certain accuracy. This scheme is used, for example, in electrochemical sinking. The majority of studies in the field of ECM theory have been performed for this scheme, because efficient methods for the design of complex-shaped TEs are of great importance.
2. Numerically controlled ECM is performed with a nonprofiling TE that moves along several coordinates simultaneously. In this scheme of ECM, one-to-one correspondence between the TE and WP surfaces is absent (the same surface can be produced using different TEs). In this case, the TE moves according to a prescribed program that enables one to obtain a desired WP surface shape. The productivity of this ECM scheme is lower, because, in contrast to the first scheme, only a fraction of the WP surface is machined at each instant of time. In contrast to the first scheme, this scheme requires the determination of, primarily, the tool path and, if necessary, the regularity of variation in the machining conditions. In the ECM of rather simple surfaces, the motion of the TE along only one coordinate is sufficient to realize the scheme (electrochemical milling, turning).
3. The combined scheme of ECM: a fraction of the WP surface is formed continuously during the machining and the rest of the surface (which is located near the TE working surface) is formed at the instant the machining is completed. This scheme is used

for electrochemical broaching of blind holes and grooves, that is, in the cases that the TE cannot freely leave the WP (TE overtravel is impossible).

For all the three schemes of ECM, the direct problems are solved in a similar way. The solutions of the inverse problem are known only for the first scheme, because for the other ECM schemes, the shape and dimensions of the machined surface do not enable one to determine, uniquely, the shape of the TE working surface.

We will consider the moving boundary problem for ECM (see Sect. 12.5.1) and its solution in the quasi-steady-state (Sect. 12.5.2) and local, one-dimensional (Sect. 12.5.3) approximations.

12.5.1 Problem with Moving Boundary

The majority of the known methods of solving *the direct and inverse problems* with moving boundaries in ECM were elaborated within the framework of the so-called model of ideal processes, ignoring the variation of the electrolyte properties in the machining zone owing to heat and gas generation and also the peculiarities of mass transfer in the diffusion boundary layer ([9] and references cited therein, [34–42], etc.). In this case, the distribution of current density over the WP surface is determined solely by the distribution of electric potential over the machining zone.

12.5.1.1 The Direct Problem

In the *nonsteady-state* case, the problem is formulated in the region with moving boundary as follows [38]:

$$\begin{aligned} &\operatorname{div}(\operatorname{grad}\varphi) = 0 \text{ in } \Omega^t \\ &\frac{\partial \Phi}{\partial t} - \frac{\eta\varepsilon\chi}{\rho}\operatorname{grad}\varphi\cdot\operatorname{grad}\Phi = 0 \text{ on } A^t \end{aligned} \tag{11}$$

where φ is the electrical potential, for Φ see Table 1.

For simple-shaped electrodes (plane, cylinder, and sphere), the problem can be solved analytically [1–9]. For more complex-shaped electrodes, numerical methods are used. For the numerical solution of the set of equations (11), both iteration methods (Ref. 36, for example) and the method of variation of inequalities are used. The latter allows one to reduce the initial problem to an equivalent problem in the region with fixed boundaries [37, 38].

In the *steady-state case*, the mathematical description of the ideal ECM process is reduced to the solution of Laplace's equation for the electric potential in the region with an unknown free boundary:

$$\begin{aligned} &\operatorname{div}(\operatorname{grad}\varphi) = 0 \text{ in } \Omega^t \\ &\operatorname{grad}\varphi\cdot\operatorname{grad}\Phi = \frac{\eta\varepsilon\chi}{V_c\rho} \text{ on } A^t \\ &g\left(\varphi, \frac{\partial\varphi}{\partial \boldsymbol{n}}\right) = 0 \text{ on } A^t \end{aligned} \tag{12}$$

where g is a function representing the boundary conditions; V_c is the TE feed rate.

Equations (12) can be solved by introducing new variables; by using them, the unknown region is transformed into a known one. For sufficiently simple shapes of the TE working surface, analytical solutions can be derived [42]. Otherwise, the problem can be solved by numerical methods [36]. Figure 8 gives the solution of the direct ECM problem for machining with a TE having vertical side surfaces at various equilibrium IEGs, S_0 (S_0 is the gap, which is reached under the conditions of constant voltage U and V_c):

$$S_0 = \frac{\eta\varepsilon\chi(U - E_a + E_c)}{V_c\rho} \tag{13}$$

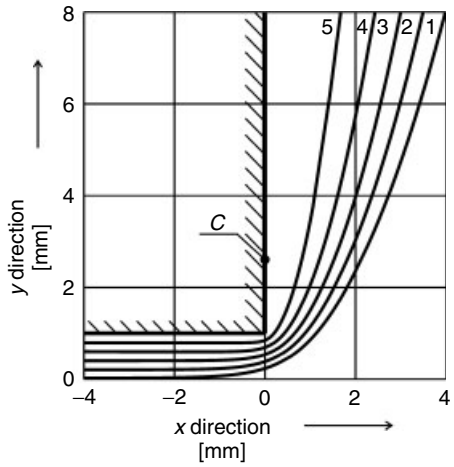


Fig. 8 The results of solving the direct ECM problem: *C* is the TE with vertical side surfaces; 1–5 are the machined WP surfaces at various equilibrium interelectrode gaps S_0 (mm): (1) 1, (2) 0.8, (3) 0.6, (4) 0.4, and (5) 0.2.

The variation in S_0 leads to the variation in the shape of the machined surface (Fig. 8). Consequently, control of machining parameters enables one to change the shape and dimensions of the machined surface. It is seen from (13) and Fig. 8 that decreasing the voltage and raising the TE feed rate allows one to decrease S_0 and improve the reproduction of TE shapes on the WP.

12.5.1.2 The Inverse Problem

All known solutions for the *inverse ECM problem* were obtained only for the steady-state case, when the problem is reduced to the solution of the Cauchy problem for the Laplace's equation for electric potential.

For the two-dimensional case, using the methods of the theory of complex variables, analytical equations for determining the shape of the TE working surface were derived [34, 39]. Figure 9 gives the analytical solution for the inverse problem. It is seen that the desired WP surface (*A*) can be obtained by using various TEs (1, 2, 3, or 4), depending on the machining parameters.

The analytical calculations involve determining the auxiliary functions of complex variables. Except for the simple shapes of WP surfaces, this presents some difficulties. Therefore, along with the analytical methods of solving the inverse ECM problem, numerical methods have evolved. Normally, in the numerical solution, the variables are changed in order to reduce the initial problem to a problem in the region with the known boundary.

The inverse ECM problem is ill posed, and its solution exists only for the sufficiently smooth WP surface. For irregular WP surfaces (which cannot be defined analytically), no analytical solution exists. Therefore, the exact solution of the inverse problem involves great difficulties and approximate methods have been developed. In these methods, the TE's shape and dimensions are determined for the precision machining of the WP surface at a number of prescribed points [36]. Between these points, the machined surface can differ slightly from the prescribed one. This approach considerably simplifies solving the inverse ECM problem.

Though the methods of solving the ECM problem with a moving boundary

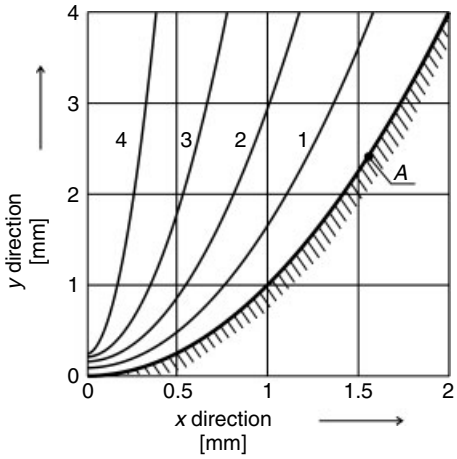


Fig. 9 The results of solving the inverse ECM problem: *A* is the desired WP surface given by the equation $y = x^2$; 1–4 are various TE surfaces that are used to produce the desired WP surface at the corresponding equilibrium interelectrode gaps S_0 (mm): (1) 0.1, (2) 0.2, (3) 0.3, and (4) 0.4.

are of great importance, they are of limited practical utility for two reasons. First, the exact solutions can be obtained only for two-dimensional problems and for sufficiently simple shapes of the TE and WP. Second, the conditions of ideal ECM are usually not fulfilled, because the electrolyte properties change significantly when the intense transfer processes proceed in the IEG. Therefore, the exact solution, which is obtained within the framework of an ideal ECM process, may differ considerably from the actual result of machining.

Though, in recent years, the solutions of the ECM problems with moving boundaries, which account for the dependence of current efficiency on the current density, have been obtained [42], it is difficult to solve the ECM problem with a moving boundary taking into account the physicochemical processes on the electrodes and in the IEG. Therefore, the approximate methods reducing the initial problem to the problem in the region with fixed boundaries were developed. The most popular approximate methods are the quasi-steady-state and local, one-dimensional approximation methods.

In the general case, the TE and WP surfaces differ substantially. The determination of the ECM rate field requires consideration of the transfer processes in the IEG, which is a complex-shaped, three-dimensional channel. In this case, the quasi-steady-state approximation method is employed.

In cases where the electrode surfaces differ insignificantly and a one-to-one correspondence between them can be reached, the hydrodynamic velocity components normal to the electrode surfaces are negligibly small and the electrical field in the IEG is quasi-homogeneous (except for the near-electrode layers). In this case, the local, one-dimensional approximation method is used.

12.5.2 Quasi-steady-state Approximation

The method of quasi-steady-state approximation (the step method) implies the use of the explicit difference scheme for solving the problem with moving boundaries. In doing so, the ECM rate field at the instant time t is calculated from the transfer equations (8) in the region Ω^t (Fig. 7) with

the known boundary Γ^t. Then, using a difference scheme, which is explicit with respect to V_a, a new WP surface shape at the instant time $t + \Delta t$ is determined:

$$\Phi^{t+\Delta t} = \Phi^t + \Delta t V_a^t |\mathrm{grad}\Phi| \qquad (14)$$

As a result, the solution of the problem with a moving boundary is reduced to the solution of a series of problems in the region with a fixed boundary.

The quasi-steady-state approximation may be used, because the rates of the transfer processes in the IEG (meters per second) are considerably higher than the rate of the variation of the WP surface (millimeters per minute). Within the framework of the quasi-steady-state approximation, it is possible to divide the initial problem into two subproblems: (1) Calculation of the transfer processes in the IEG and the determination of the ECM rate field V_a. (2) Calculation of the WP surface evolution for the direct problem or correction of the TE surface for the inverse problem. However, even under this simplification, solving the direct and inverse ECM problems, especially for sculptured WPs, involves great difficulties.

12.5.2.1 Calculation of Transfer Processes in the IEG

The transfer processes in the gap within the quasi-steady-state approximation are calculated similarly for both the direct and inverse problems. To simplify the calculation of transfer processes in the gap, the boundary-layer approximation is commonly used. According to this approximation, the current density is calculated separately in the bulk gap and in the near-electrode diffusion layers, and their congruence is provided via the boundary conditions. The transfer processes in the near-electrode diffusion layers are often taken into consideration using the experimental polarization curves.

12.5.2.1.1 Calculation of Transfer Processes in the Bulk IEG

In the quasi-steady-state approximation, the distribution of the current density over the bulk IEG is described by Laplace's equation:

$$\mathrm{div}(\chi\ \mathrm{grad}\ \varphi) = 0$$
$$\boldsymbol{i} = -\chi\ \mathrm{grad}\ \varphi \qquad (15)$$

For solving Eq. (15), appropriate boundary conditions must be prescribed. Normally, the boundary of the machining zone consists of several sections at which the boundary conditions of different types are prescribed. The type of the boundary conditions depends on the character of the boundary section: TE, WP, insulator, or the line (the plane of symmetry), and also on the operating conditions of the power supply: the conditions of stabilization of the applied voltage, the conditions of current stabilization, and the natural current–voltage characteristics. In the general case, the boundary conditions that account for the kinetics of the electrode reactions and the transfer processes in the near-electrode diffusion layers can be written as follows:

$$g\left(\varphi, \frac{\partial\varphi}{\partial\boldsymbol{n}}\right) = 0 \qquad (16)$$

Various numerical methods are used to solve Laplace's equation for ECM including the method of finite differences, the finite element method, the boundary element method [9, 43, 44], and so forth.

Since the regularities of gas transfer from the electrode surface to the bulk electrolyte are not clearly understood, it is commonly assumed that $\chi = \mathrm{const}$

[45–49]. In this case, the boundary element method is most suitable. Laplace's equation is solved in the following way (Fig. 10).

1. Three-dimensional, solid-state models of the TE and WP are developed (Fig. 10a).
2. The machining zone is formed. The boundary of this zone consists of the TE (3), the WP (4), the insulator surfaces, and the planes of symmetry (5). The surface of this zone is meshed by triangle or quadrilateral boundary elements (Fig. 10b). In each element, the potential and its gradient are taken to be constant.
3. For the given scheme of partition of the machining zone boundary on the elements, the discretization of the boundary integral equation and the boundary conditions is performed. The set of nonlinear equations, which is obtained by discretization, is solved by Newton's method. As a result, the distribution of the current density over the WP surface is obtained (Fig. 10c).

In some works [50, 51], the variations in the electrolyte properties due to the heat and gas evolution in the gap were taken into consideration. It was assumed that the gas phase was uniformly distributed across the gap. The consideration of variation in the electrolyte properties leads to the necessity of numerical solution of equations describing the flow of viscous anisothermic gas–liquid mixtures simultaneously with Laplace's equation.

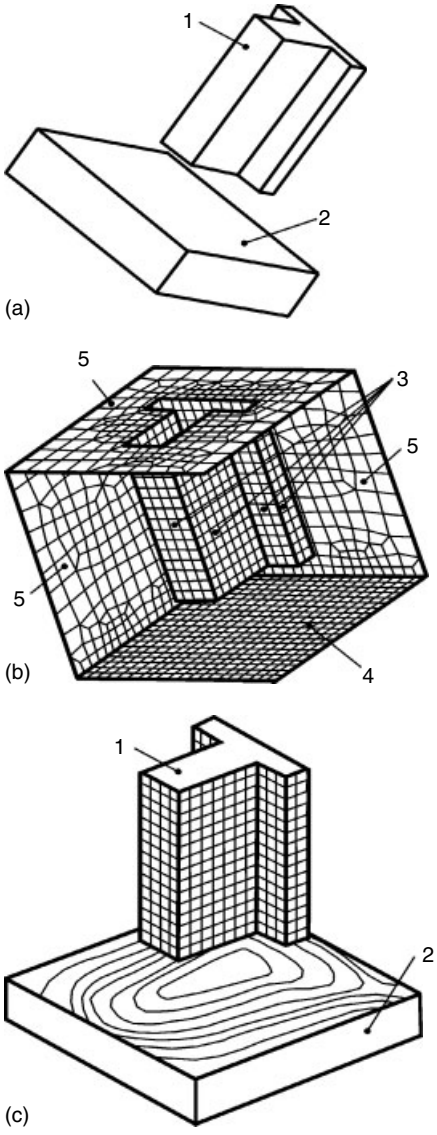


Fig. 10 Calculation of the current density distribution over the bulk of interelectrode space within the quasi-steady-state approximation using the boundary element method: (a) Three-dimensional solid-state models of TE and WP, (b) scheme of partition of machining zone boundary on the quadrilateral boundary elements (by convention, two side surfaces are not shown), (c) the current density distribution (the current density level lines) over the WP surface; (1) the TE, (2) the WP, (3, 4) the working surfaces of the electrodes, and (5) the planes of symmetry and the insulator.

12.5.2.1.2 **Calculation of Transfer Processes in the Near-electrode Diffusion Layer**

The calculation of the transfer processes in the near-electrode diffusion layer is based on the set of equations of anisothermic ionic mass transfer, which is caused by diffusion, migration, convection, and homogeneous chemical reactions:

$$\frac{\partial C_k}{\partial t} = -\mathrm{div}(\boldsymbol{j}_k + C_k(\hat{V} + \mathrm{grad}\ \Psi)) + R_k \qquad k = 1, \ldots, N$$

$$\bar{\rho}\mathrm{div}(\mathrm{grad}\ \Psi) = \sum_{k=1}^{N} \mathrm{div}(\hat{M}_k \boldsymbol{j}_k) - \sum_{k=1}^{N} \hat{M}_k R_k$$

$$\frac{\partial Q}{\partial t} + \mathrm{div}(Q[\hat{V} + \mathrm{grad}\ \Psi]) = -\boldsymbol{i}\ \mathrm{grad}\ \varphi + \mathrm{div}(\lambda\ \mathrm{grad}\ T)$$

$$\sum_{k=1}^{N} z_k C_k = 0 \qquad (17)$$

where C_k is the concentration of the kth electrolyte component; $\mathbf{j}_k = -D_k\ \mathrm{grad}\ C_k - z_k u_k F C_k\ \mathrm{grad}\ \varphi$ is the diffusion flux of the kth electrolyte component; D_k, z_k, u_k are the diffusion coefficient, the charge, and the mobility of the kth electrolyte component, respectively; $\hat{\mathbf{V}}$ is the hydrodynamic velocity, which is determined by external factors; Ψ is the potential of hydrodynamic velocity, induced by intense interface mass transfer; R_k is the intensity of the internal source of the kth electrolyte component, which is associated with the chemical reactions; $\bar{\rho}$ is the solvent density; $\hat{M}_k$ is the effective molecular mass of the kth electrolyte component; $Q = c_p \rho T$ is the specific heat energy; c_p is the specific heat; $\mathbf{i} = F\sum_{k=1}^{N} z_k \mathbf{j}_k$ is the current density; and λ is the thermal conductivity.

At not-too-high current densities, the electrolyte flow induced by the interface mass transfer (the second equation in system (17)) and the variation of the temperature in the diffusion layer (the third equation in system (17)) can be neglected.

For a binary electrolyte and solution with a high content of supporting electrolyte, the system of equations (17) can be solved analytically. For the electrolyte with complex composition, when migration transfer cannot be ignored, numerical methods are used. Two approaches to the numerical solution of ionic transfer equations in the ECM can be considered [52–55]. The first approach consists of simultaneously solving all equations of the system (17). The second approach involves decoupling of the initial set of ionic transfer equations, which allows the successive determination of the electrical potential, potential of velocity induced by the interface transfer, concentrations of components, and electrolyte temperature. The second approach significantly reduces the computational time, enabling one to perform the calculations not only for the continuous ECM but also for pulse and pulse-cyclic ECM.

To determine the field of ECM rate, the transfer processes are successively calculated in the bulk IEG and in the diffusion layers, until a consistent solution is obtained. Using the obtained field of ECM rate, new shapes and dimensions of the WP surface are determined.

12.5.2.2 **Direct Problem: Evolution of WP Surface**

In considering the evolution of the WP surface in the quasi-steady-state approximation, methods based on the explicit coordinate and parametrical description of the WP surface are most popular. These methods are described in detail in Refs 43, 56, 57. The explicit coordinate description is commonly used for sufficiently simple geometrical shapes of the WP surface,

which can be represented as a graph. The parametrical presentation allows one to extend the range of surfaces; however, it requires the use of special procedures for adequately simulating the surface evolution in the areas where the WP surface is not smooth and in the zones where the WP and the insulator surfaces join [43]. The use of implicit coordinate presentations of WP surface is promising, because it makes possible simulation of arbitrary surfaces, including rough ones [58]. The possibilities of this presentation can be demonstrated using the following model of WP surface evolution (see Table 1):

$$|\text{grad } T| = \frac{1}{V_a} \tag{18}$$

where T is a function of $\boldsymbol{x}$, which is equal to the arrival time of the WP surface to a considered point $\boldsymbol{x}$ of the space.

In this case, the WP surface, A^t, is the level set of function $T(\boldsymbol{x})$, that is,

$$A^t = \{\forall x | T(x) = t\} \tag{19}$$

The determination of the location of the WP surface at every instant of time involves the solution of Cauchy's problem for Eq. (18) with the initial condition being given in the following form:

$$T(x) = 0, \forall x \in A^0 \tag{20}$$

where A^0 is the initial WP surface.

For smooth surfaces having topology invariant to the machining, the solution of equations (18–20) is relatively simple. However, often the conditions of smoothness and the absence of topological variations are not fulfilled. In the ECM of even initially smooth surfaces, surface irregularities can form on the WP in the course of machining, and it is impossible to predetermine their location and the instance of their rise. Equation (18) admits several solutions, but only one of them can be physically realized [48, 58, 59]. Figure 11 gives the results of simulating the evolution of the initially plane WP surface in the ECM using a rough T-shaped TE, which is shown in Fig. 10.

12.5.2.3 The Inverse Problem: Correction of TE Surface

Within the quasi-steady-state approximation, the most popular method is the "embedding method" involving the successive solving of a series of direct problems and corrections of the TE working surface until a desired shape of the WP is assured [45, 46, 50, 51]. Several methods of TE correction have been proposed; they have a lot in common and can be presented in the following generalized form:

1. For the desired WP surface $z_a(x, y)$, the TE working surface can be presented as a series

$$z_c = \sum_{k=1}^{K} a_k f_k(x, y) \tag{21}$$

 where $f_k(x, y)$ is a family of functions used to represent the TE; a_k are the coefficients of the expansion; and K is the number of coefficients.
 At the first iteration, the magnitudes of a_k are prescribed, that is, the initial approximation of the TE working surface is described.
2. The direct ECM problem is solved, that is, the real WP shape, $z_a^*(x, y)$, which will be obtained as a result of machining with a given TE, is determined.
3. A deviation $E(\mathbf{a})$ of the actually obtained WP shape from the desired one is determined ($\mathbf{a}$ is a vector of the TE shape coefficients a_k). A mean square root or maximum deviation of $z_a^*(x, y)$ from $z_a(x, y)$ can be used as $E(\mathbf{a})$. If $E(\mathbf{a})$

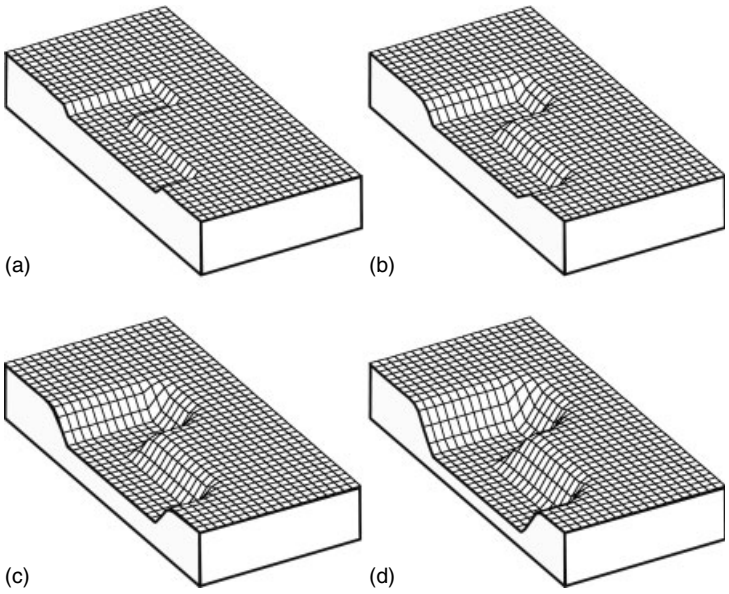


Fig. 11 Variation in the WP shape during ECM. The time after the machining was started (min): (a) 1, (b) 2, (c) 3, (d) 4.

exceeds the admissible error, the TE is corrected; otherwise, the calculations are finished.

4. $E(\mathbf{a})$ depends only on the vector of the TE shape coefficients; therefore, the TE correction is reduced to the minimization of $E(\mathbf{a})$, that is, to the solution of the following set of equations:

$$\frac{\partial E(\mathbf{a})}{\partial a_k} \approx \frac{E(\mathbf{a} + \Delta a_k) - E(\mathbf{a})}{\Delta a_k} = 0 \tag{22}$$

Here Δa_k is a small perturbation imposed on coefficient a_k.

The derivatives in Eq. (22) are calculated numerically; this requires K direct model evaluations. To solve the set of nonlinear equations (22), it is advantageous to use the Newton–Raphson method possessing quadratic rates of convergence and requiring no more than four to seven iterations [45].

12.5.3 Local, One-dimensional Approximation

Within the local, one-dimensional approximation, the solution of the direct and inverse ECM problems is considerably simplified.

First, instead of determining the shape of one electrode, it is sufficient to find the distribution of the IEG over the machining zone. This appears to be much simpler than determining, immediately, the electrode's shape, because the IEG depends only on three parameters (u, v, t), whereas the electrode surface is determined by four parameters (x, y, z, t).

Second, using the local, one-dimensional approximation for the distribution of the electrical potential, one can obtain the following approximate equation for the gradient of electrical potential:

$$\operatorname{grad} \varphi \approx \frac{\partial \varphi}{\partial \varsigma} = \frac{U - E_a + E_c}{S} \tag{23}$$

where ς is the coordinate measured from the electrode surface along the outward normal.

Third, in the system of coordinates with the origin located at the surface of one of the electrodes, the hydrodynamic velocity field is two dimensional. Therefore, prescribing the distributions of hydrodynamic velocity, gas fraction, temperature, and so forth across the IEG, one can integrate the equations of mass, momentum, and energy transfer with respect to the distance between the electrodes. As a result, it is possible to reduce the problem's dimension by a unit.

As well as in the quasi-steady-state approximation, within the local, one-dimensional approximation, the solution of the direct and inverse ECM problems is reduced to successively solving the following problems: (1) Calculation of the transfer processes in the IEG and determination of the ECM rate field V_a. (2) Calculation of the IEG distribution. The calculated IEG distribution is used to determine the WP surface in the direct problem and to correct the TE surface in the inverse problem.

In the ECM, the electrolyte temperature and the void fraction increase in the direction of electrolyte flow, leading to a nonuniform distribution of current density over the WP surface. With a decrease in the IEG width, the degree of current density nonuniformity and, hence, the ECM inaccuracy, increase (Fig. 12). An adverse effect of heat and gas evolution can be reduced significantly by passing to the pulse ECM mode. At sufficiently short pulse-on times, when electrolyte convection can be ignored, the current density is independent of the IEG length in the flow direction and is determined solely by the IEG width. In the initial period of time after imposing a pulse, the current density is higher for a smaller gap (Fig. 13). Therewith, the localization of anodic dissolution is high, the dissolution is localized on the areas with the smallest gap, and the initial WP inaccuracy decreases sharply. At $t = t_1$, the current density on the areas with different gaps is leveled, and at $t > t_1$, the ECM inaccuracy starts increasing. At $t > t_2$, when the charge passed on the areas with a smaller gap becomes smaller than that on the areas with a larger gap, the accuracy of pulse ECM becomes even lower than the accuracy of constant-voltage ECM. In many cases of pulse ECM, it is

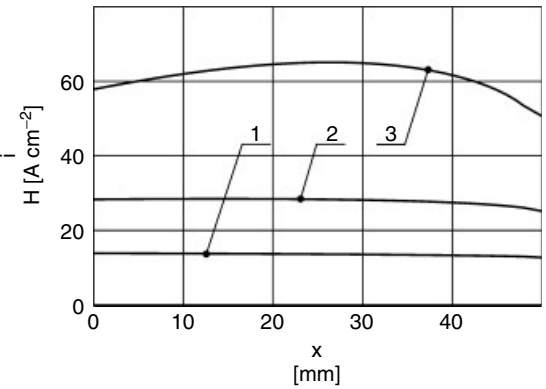


Fig. 12 Distributions of the current density along the WP length (in the direction of electrolyte flow) in the ECM with a constant voltage at various interelectrode gaps (mm): (1) 1, (2) 0.5, and (3) 0.25; $x = 0$ is the electrolyte inlet to the IEG, $x = 50$ is the electrolyte outlet.

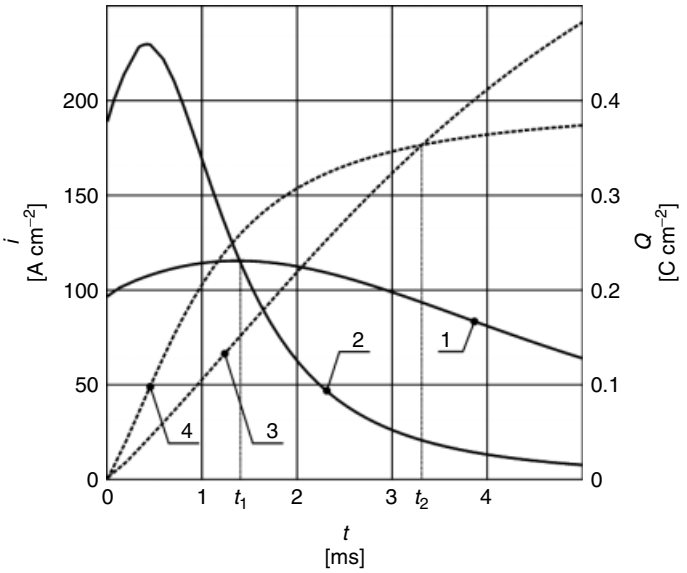


Fig. 13 Time variation in the (1, 2) current density and (3, 4) charge passed in pulse ECM at various interelectrode gaps (mm): (1, 3) 0.1 and (2, 4) 0.05.

advantageous to choose a pulse-on time equal to t_1: in this case, a maximum decrease in the initial inaccuracy in a single pulse is observed. At shorter pulse-on times, the degree of ECM localization increases, but the productivity significantly decreases.

12.5.3.1 Calculation of Transfer Processes in the IEG

To calculate the transfer processes in the IEG, the boundary-layer approximation is used. According to it, the current density in the bulk IEG is calculated on the basis of the height-average mass-, momentum-, and energy-transfer equations and those in the near-electrode layers. The transfer in the diffusion layers is calculated similarly to the case of quasi-steady-state approximation.

The distribution of the current density over the bulk IEG is calculated using the averaged-over-the-IEG Navier–Stokes equations for homogeneous, gas–liquid mixtures. For the TE, which is given parametrically by the vector function

$$\boldsymbol{r}^{\mathrm{c}} = \boldsymbol{r}^{\mathrm{c}}(\xi, \vartheta), \quad \boldsymbol{r}^{\mathrm{c}} = (x^{\mathrm{c}}, y^{\mathrm{c}}, z^{\mathrm{c}}) \qquad (24)$$

in the curvilinear orthogonal coordinate system $(\xi \vartheta \varsigma)$ with the origin at the TE surface, the Navier–Stokes equations averaged over the IEG, which describe two-dimensional flow of gas–liquid mixture in the IEG, are commonly presented in the matrix form [60]:

$$\frac{\partial \boldsymbol{W}}{\partial t} + \frac{\partial \boldsymbol{F}}{\partial \xi} + \frac{\partial \boldsymbol{G}}{\partial \vartheta} = \boldsymbol{B} \qquad (25)$$

Here $\boldsymbol{W}$ is a vector with the elements being the conservative variables: the gas phase and mixture densities, the momentum, and the energy of the mixture; $\boldsymbol{F}$ and $\boldsymbol{G}$ are the fluxes of vector $\boldsymbol{W}$ elements in

the direction of coordinate axis ξ and ϑ, respectively; $\boldsymbol{B}$ is the source term.

Equation(25) involves the continuity equations for the gas phase and the mixture, two mixture momentum balance equations for the coordinate directions ($\xi\vartheta$), and the energy equation.

To numerically solve Eq. (25), the finite volume method is used. In doing so, a mesh consisting of triangle cells is constructed at the TE surface; in each cell, the flux parameters are taken to be constant. The fluxes through the boundaries of finite volumes are determined on the basis of the solution of the one-dimensional Riemann problem in the direction of normal to the boundary between cells.

As a result, the fields of hydrodynamic velocity and pressure and the distributions of temperature and void fraction are determined. The current density is determined by the following equation:

$$i = \frac{\chi(1-\alpha)^n[1+\beta(T-T_0)](U-E_a+E_c)}{S} \tag{26}$$

where $(1-\alpha)^n$ is the term that takes account of the effect of void fraction on the electrolyte conductivity; $\beta(T-T_0)$ is the term that accounts for the effect of temperature on the electrolyte conductivity.

Figure 14 gives the results of simulation of transfer processes in the bulk of the interelectrode space in the quasi-equidistant ECM. Using a CAD system, a three-dimensional solid-state or surface TE model is formed (Fig. 14a). A mesh of triangular finite volumes is constructed at the TE working surface (Fig. 14b). The field of velocities of electrolyte flow in the gap (Fig. 14c) and the distribution of the current density (Fig. 14d) are then calculated.

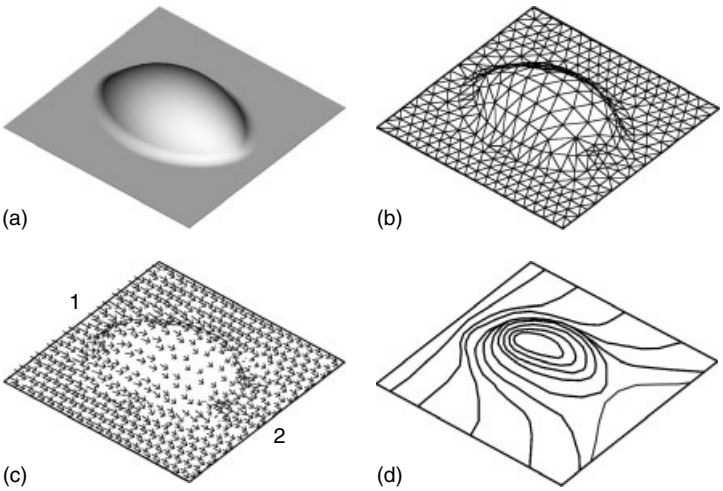


Fig. 14 Calculation of transfer processes in the bulk of the interelectrode space within the local, one-dimensional approximation: (a) Solid-state model of the TE, (b) scheme of partitioning of the working surface of the TE with triangular elements, (c) electrolyte flow velocity field, (d) distribution of the current density over the TE surface; (1) is the electrolyte inlet and (2) is the electrolyte outlet.

12.5.3.2 The Direct Problem: The Evolution of WP Surface

Much research has been devoted to simulating the evolution of the WP surface in the quasi-equidistant ECM [1–9, 42, 61, 62]. In most cases, models that allow one to perform the calculations not over the entire surface but only in several plane sections of the IEG are used. The use of the theory of surfaces allows one to obtain the generalized model for calculating the IEG distribution as [63]:

$$\frac{\partial S}{\partial t} = V_a\sqrt{1+\Delta_1 S_{|a}} - \boldsymbol{V}_c \cdot \boldsymbol{n}^c + \mathrm{div}(S\boldsymbol{V}_c)_{|a} \tag{27}$$

where $\boldsymbol{V}_c$ is the vector of the TE feed rate and $\Delta_1 S_{|a}$ is the Beltrami first differential parameter of the IEG on the WP surface.

There are three terms in the right-hand side of Eq. (27): the first term allows for the variation of the IEG due to the WP anodic dissolution; the second term allows for the variation of the IEG that is attributed due to the normal to the TE surface component of the feed rate; and the third term allows for the variation of the IEG due to the WP surface motion along the TE surface. Using Eq. (27), the variation in the IEG magnitude can be determined for any scheme of quasi-equidistant ECM.

The numerical solution of Eq. (27) is performed by the finite volume method [63]. A mesh of triangular elements is constructed on the working surface of the TE. It is assumed that the curvature of the TE surface is constant within each element. The principal curvatures and principal directions are determined in the center of the element, when the mesh is constructed. The calculations are performed in the local system of coordinates with the origin in the center of the triangular element; the u and v axes coincide with the directions of principal curvatures. The set of ordinary differential equations, which is obtained as a result of discretization, is integrated by the Runge–Kutta method.

Simulation of the WP surface evolution is performed as follows: (1) Construction of a solid-state TE model; (2) transformation of the solid-state model into the surface model; (3) partition of the TE surface into triangular elements; (4) calculation of IEG distribution; and (5) calculation of WP surface using Eq. (4).

12.5.3.3 The Inverse Problem: The Correction of TE Surface

To solve the inverse ECM problem in the local, one-dimensional approximation, the above-described method for the quasi-steady-state approximation can be used (see Sect. 12.5.2.3). However, especially for complex-shaped surfaces, it is advantageous to use a simpler and more efficient method. When correcting the TE surface, the error in the WP surface shape is conveniently estimated along the normal to the TE surface. This, on the one hand, does not lead to considerable error, because, in the quasi-equidistant ECM, the normals to the TE and WP surfaces differ little, and, on the other hand, enable one to significantly simplify the correction procedure. The TE correction is realized as follows:

1. For a desired WP surface, the initial approximation of the TE working surface is prescribed and the natural parameters of the TE surface and their relation to the natural parameters of the machined surface are determined.
2. For the chosen machining conditions, the distribution of gap width over the machining zone is calculated using Eq. (24).
3. The inaccuracy of the machined surface is determined. If the inaccuracy

exceeds the permissible value, the TE is corrected; otherwise, the calculations are terminated.

4. The TE working surface is corrected via shifting each computational node, which is located on the TE surface, in the direction of the normal to the TE surface by a value equal to a difference between the required and actual gap magnitudes in this node. Then, the natural parameters of the corrected TE and their relation to the WP surface parameters, and the required distribution of gap width for the corrected TE, are calculated.

In contrast to the above-described method (see Sect. 12.5.2.3), this method requires only a single calculation of the distribution of current density at each iteration and has a comparable convergence rate.

12.5.4 Automated Design of ECM Operations

A typical scheme for the design of ECM operations involves the solution of the following problems: choosing the machining scheme, the electrolyte composition, the equipment, and machining conditions; the determination of allowance and the design of necessary machining attachments; testing; and correction.

In the design, the optimal conditions of machining, which provide high productivity of machining, the required accuracy, and high surface finish, should be determined.

Some problems, such as choosing the machining scheme, the electrolyte composition, the machining equipment and attachments, can be solved using typical solutions. The main problems in the design of ECM operations are associated with the design of the TE and the determination of the machining parameters. The accuracy and productivity of finishing operations depend to a large extent on the proper solution of these problems.

The difficulty in the design of ECM operations is that the machining parameters and the shape of the TE working surface cannot be determined independently of one another. When the machining parameters are changed, the IEG distribution also changes; consequently, to produce a WP of the desired shape and size, the working surface of the TE should be corrected.

Figure 15 gives the general scheme of an automated design of ECM operations. Using an integrated CAD/CAM/CAE system, the three-dimensional, solid-state or surface model of the WP is created. To elaborate an ECM operation, a geometrical model of the WP and also other required information (the WP material, the accuracy of dimensions and shape, and the surface finish) are delivered from the universal CAD system to the subsystem of ECM operation design.

The subsystem for the automatic design of ECM operations in the dialog mode enables the following problems to be solved: (1) Analysis (the direct problem): the shape and size of the WP surface are determined by the given TE surface; and (2) synthesis (the inverse problem): the working surface of the TE is determined in order to produce the desired WP surface.

The use of the dialog mode in the design of ECM operations is caused by the fact that the mathematical models of ECM are rather complicated, and the solution of the problem depends on a large number of parameters. Realization of the dialog mode of design actually means the replacement of the machining of test WPs (the trial

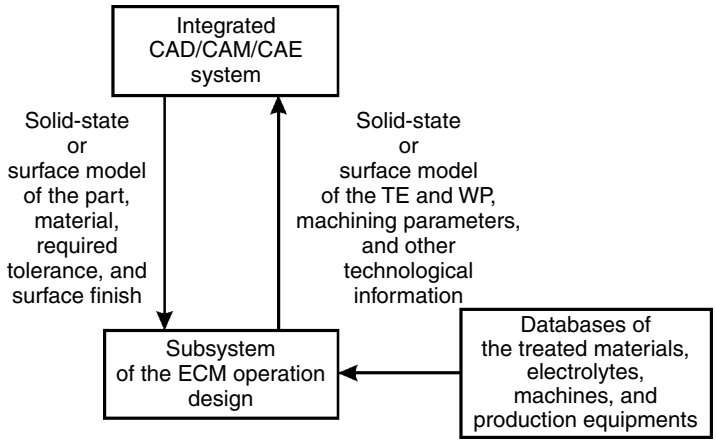


Fig. 15 General scheme of automated design of ECM operations.

and error method) by the "machining" of geometrical models of the test WPs in the course of computing experiments.

The solution of the inverse problem is of principal commercial importance. However, the exact solution of this problem is difficult, because it requires the solution of a problem with unknown boundaries. Therefore, for the automation of the design of ECM operations, the following approach is used: the inverse problem is reduced to the solution of a series of direct problems and, by solving them, the TE working surface and the machining parameters are corrected.

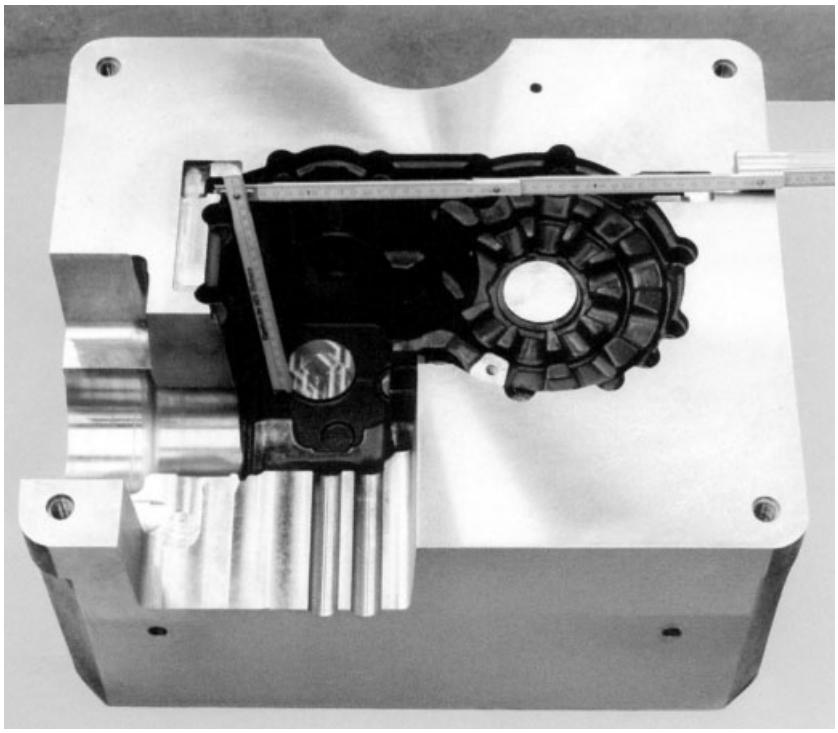
The output information of the subsystem of automated design of ECM operation involves geometrical models for the WP and TE, the machining parameters, and other technological information. This information is delivered to the integrated CAD/CAM/CAE system for the formation of technology documentation, development of technology of attachment and TE fabrication, and specification of programs for machines with digital control for producing the TE working surfaces.

12.6 ECM Techniques and Applications

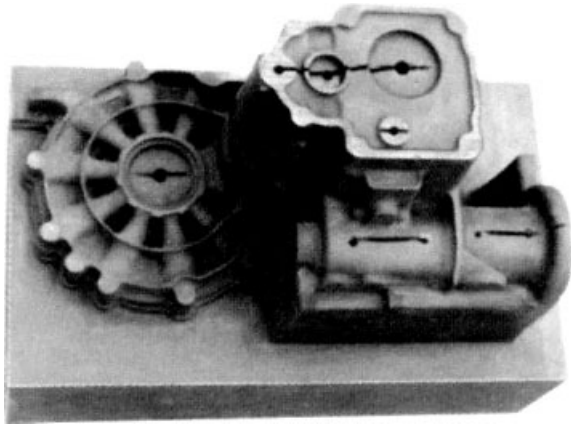
12.6.1 Electrochemical Sinking

ECM sinking is effectively used for the production of turbine and compressor blades, dies, molds, and other complex-shaped products, which are made of hard-to-cut steels and alloys. In ECM sinking, in contrast to the traditional methods of machining, the entire WP surface is machined in a single operation. For example, ECM sinking was used for machining large dies of hardened and tempered steel for the automotive industry (Fig. 16) [64]. Application of ECM enabled one to obtain the required accuracy and surface finish in a single operation; the production time was cut by more than four times compared with electrodischarge machining (EDM).

In ECM sinking, the accuracy and surface finish depend considerably on the proper choice of TE design, a scheme of electrolyte feed through the slots and holes into the machining zone, the scheme, and parameters of machining. ECM sinking is



(a)



(b)

Fig. 16 (a) Die produced by ECM sinking and (b) TE. Electrolyte feed holes and slots are seen on the TE surface.

realized by various schemes: with constant TE feed rate, with TE vibration, and by the cyclic scheme.

In order to realize ECM sinking, a new generation of electrochemical machines, equipped with adaptable control systems and wide-range pulse power sources, has been developed in the last decade. Figure 17 gives an ECM-1500A machine (Novotech) for pulse ECM with vibration



Fig. 17 ECM-1500A machine (Novotech Ltd) for electrochemical three-dimensional shaping.

of TE [65]. The machine is equipped with an automatic control system that provides effective adaptation of the processing to the variations of external conditions and guarantees stability of accuracy and roughness. The short-circuit protection system contains three basic contours: the contour for prediction of short circuit or breakdown in the interelectrode space; the contour for current switch-off in the case of short circuit; and the contour for protection of power source against overloads. An ECM-1500A machine provides an accuracy of dimensions of $\pm 5\,\mu m$ and copying resolution of $\sim 1\,\mu m$ at a surface roughness (R_a) of 0.05 to 0.2 µm. The machining is performed at an IEG of 3–10 µm and current densities up to 1000 A cm^{-2}. Figure 18 gives a specimen of 40Cr13 steel, on which the rows of cylindrical and rectangular projections are produced with an ECM-1500A machine; the error of each element does not exceed 5–7 µm, which is comparable with the accuracy of EDM. However, the absence of TE wear, the defect layers on the machined surface, and a higher productivity are the obvious advantages of ECM.

12.6.2 Electrochemical Drilling

Electrochemical drilling (ECD) is an important application of ECM. The ECD operations enable the production of not only round holes, as in the case of mechanical drilling, but also shaped holes with oval, rectangular, triangular, and other sections; to drill several tens of holes of varying diameters and/or shapes simultaneously; to produce curved holes or slots, for example, ports and passages in manifolds; to produce holes that are not perpendicular to the surface, and to

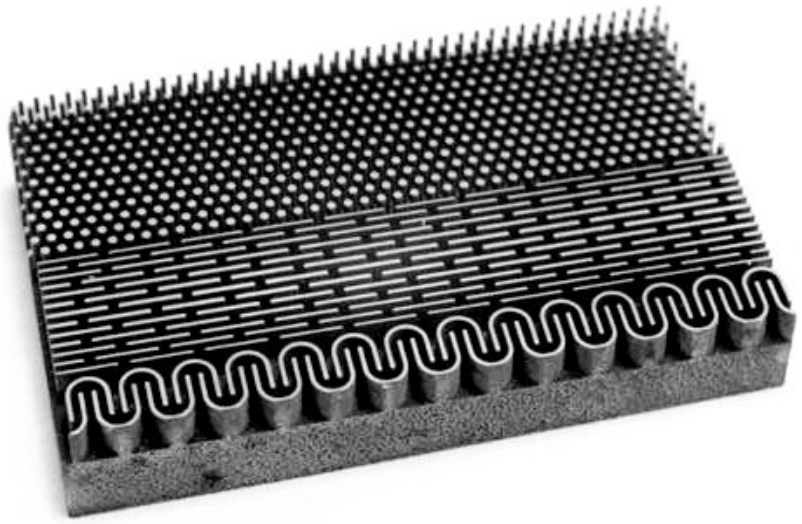


Fig. 18 Demonstrative specimen of 40Cr13 steel (20 × 30 mm) machined with ECM-1500A machine.

perform other unique drilling operations in the difficult-to-machine materials [45, 66–74]. Figures 19–21 give schemes of drilling of curved hole and a hole with a contoured surface (a hole with internal ridges, which ensure turbulent airflow

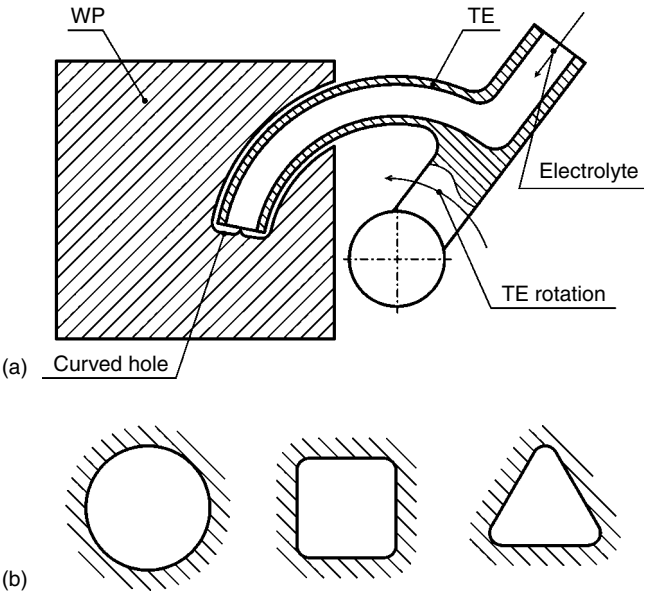


Fig. 19 (a) Scheme of curved hole ECM drilling (b) with various cross sections.

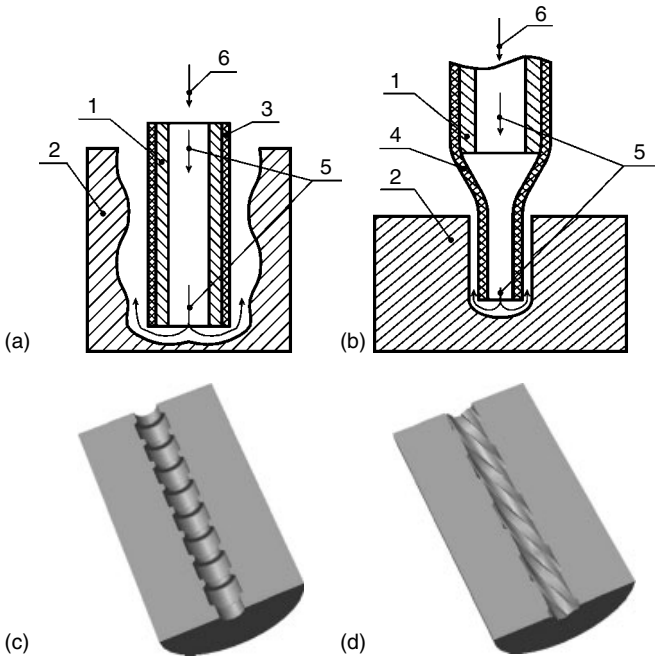


Fig. 20 (a,b) Schemes of ECM drilling and (c,d) sections of specimens, in which holes with contoured surface are produced: (1) the TE, (2) the WP, (3) the insulating cover, (4) the glass tube, (5) the electrolyte flow, and (6) the TE feed.

through the cooling holes and increase the efficiency of mass-transfer processes) using the ECM [69, 70, 74]. The contouring ECD machine is governed by a CNC machine controller to ensure proper control functions. In addition to contoured surface, these holes have small diameters (e.g. 1.3 mm) and great depths (200 mm). For this purpose, the ECD machine should be capable of high positioning accuracies and rigid control for time-dependent commands. Under these conditions, excellent repeatability of results from hole to hole can be provided.

Drilling of small deep holes with a depth-to-diameter ratio sometimes up to 100:1 and even higher in the difficult-to-machine materials (high-strength, superalloy materials) is a very important area of ECD application ([66–72]; the patent literature on ECD, as well as on other ECM problems, is quite extensive). The ECD of small deep holes was developed predominantly for producing air-cooling holes in gas turbine buckets and vanes to generate higher power output of jet engines for the same unit size and with improved fuel efficiency.

In the majority of methods of drilling of small deep holes (many of them were elaborated by the engineers of General Electric), a tube through which an electrolyte is pumped under pressure serves as the tool. The tube can be made of metal (Fig. 20a); in this case, the tube itself is a cathode. It should be somewhat smaller in diameter than the required hole. The metal tube is coated externally (except for a tip area) with a film of insulation

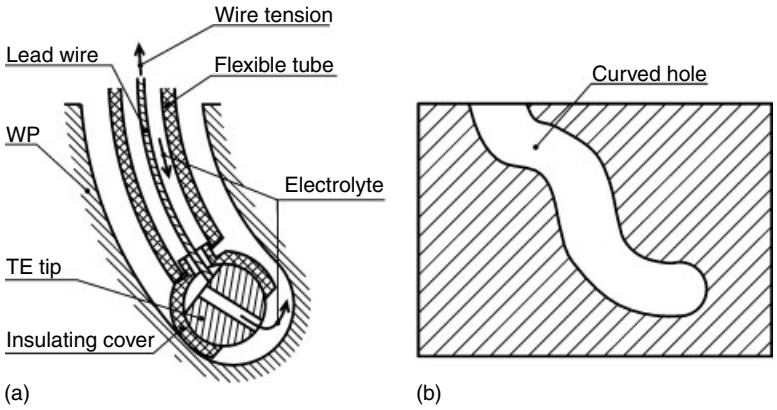


Fig. 21 (a) Scheme of curved hole ECM drilling controlled by wire tension or electrolyte pressure and (b) an example of the produced hole.

that protects the walls of the hole from dissolution. A metal tube is used when shaped holes should be drilled. In this case, the tip of the metal tube has the corresponding shape. The tube is fed toward and into the WP.

An alternate design of tool for ECD involves a glass tube (Fig. 20b). It has a significantly larger diameter than that of the hole and a relatively small-diameter opening at the tip. The outside diameter of the tip must be sufficiently smaller than the diameter of the hole to allow for the escape of the electrolyte. In this case, high pressure from the electrolyte is required for the high-rate ECD process. A metal cathode in the form of a cylinder (Fig. 20b) or wire is placed inside the thick part of the glass tube. The constant feed advances the drilling nozzle into the WP at a rate equal to the rate at which the WP material is dissolved (penetration drilling mode). The version of ECD with a stationary drilling nozzle (dwell drilling mode) is preferred when shallow, less accurate holes are required.

When a glass drilling nozzle is used, the interelectrode distance appears to be considerably larger than in the case of a metal tube. This fact, coupled with a small cross-section area of electrolyte stream, results in a large voltage drop in the IEG and the necessity of applying high voltages (of several hundreds volts). The highest voltages should be used in the case of a stationary tool.

In the ECD of fine holes, a highly conductive acid electrolyte is commonly used (10–30% nitric, sulfuric, or hydrochloric acid, depending on the WP material) in order to avoid the formation of sludge, hampering the solution flow through fine holes. When a metal tube serves as the TE and an acidic electrolyte is used, dissolved metal accumulates in the electrolyte and deposits on the cathode working surface, changing the cathode's shape. To prevent this, the voltage is periodically reversed long enough to remove film buildup on the TE but not long enough to change the prescribed TE shape.

In some cases, salt electrolytes with the addition of a complexing agent to prevent sludge formation are used [73].

The production of machinery for unique ECD operations requires a high level of capital investment.

12.6.3 Electrochemical Deburring

Electrochemical deburring (ECDB) is one of the largest applications of electrochemical metal removal technology [19, 75–79]. ECDB is a nontraditional process for removing burrs from metal parts, such burrs usually having been created by traditional machining operations. The advantages of the method are the ability to finish hard-to-reach areas, inaccessible for ordinary methods (in intersecting holes, inside of bores, or in grooves), high productivity (short machining times, deburring of several WPs simultaneously), relatively low cost, and a high repeatability of results. Figure 22 gives a simplified scheme of ECDB in intersecting holes. The major problem is to localize the dissolution solely on the burrs. Etching of surface on other WP areas, particularly adjacent to burr areas, is the main disadvantage of ECDB. To prevent this, the proper design of TE (only TE areas that will be in the immediate vicinity of burrs should not be insulated), controlled flow of electrolyte, electrolyte composition, and parameters of process are of importance. In ECDB, the TE is normally placed adjacent to the area to be machined and remains stationary during the process. Two types of electrolytes are used for ECDB: typical electrolytes for ECM ($NaNO_3$) and typical electrolytes for electrochemical polishing. In some cases, additives are used. As a rule, radiusing of sharp edges is performed simultaneously with ECDB. Components that are susceptible to corrosion must be treated immediately after the deburring operation, normally with a water rinse and then a dip in a rust inhibitor.

ECDB can be utilized in small shops and high-volume manufactures, as a manual, semiautomatic, or fully automated process. The processing step for the ECDB can be incorporated into the production line. ECDB greatly reduces deburring costs on operations which are traditionally time consuming and which require large amounts of hand labor. As a rule, a special TE is required to machine WP of each type; therefore, ECDB is most efficient in mass production. Figure 23 gives the examples of parts, which are deburred with ECM.

12.6.4 Removing of Defective Layers

Removing of defective layers from the surface of shaped parts is an important

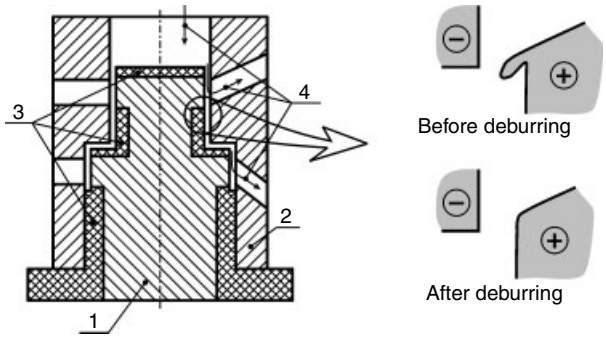


Fig. 22 Scheme of electrochemical deburring and sharp edge radiusing: (1) the TE, (2) the WP, (3) the insulating cover, and (4) the electrolyte flow.

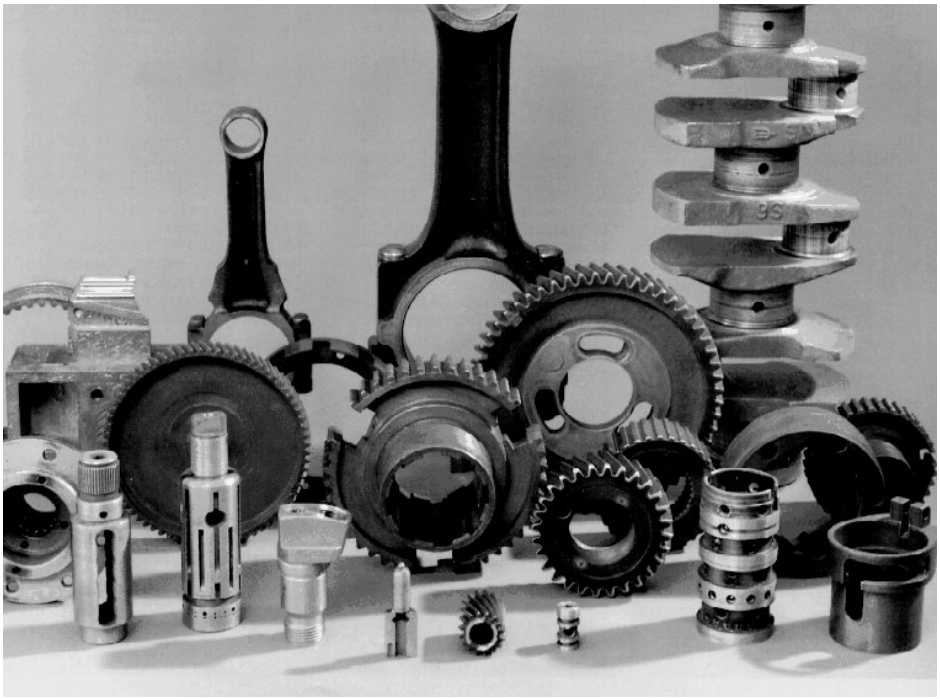


Fig. 23 Examples of parts, on which electrochemical deburring is produced (AEG-Elotherm).

application of ECM. This method is specially effective for removing of defective layers created by electrodischarge machining (EDM) [80–83]. EDM is a very important manufacturing process. However, EDM produces heat effects such as heat-affected layers, deep crater topographies, and microcracks in the surface region of the WP (e.g. dies and molds). The electrochemical process does not destroy the original shape of the WP, but it removes a damaged surface layer and surface roughness. In some cases, the surface of the cavities with various shapes can be finished in a mirrorlike and reproducible manner within a few seconds. An optimal duration of the process should be chosen in order to obtain best results (high surface finish). Usually the same machine and TE, which were used for EDM, are applied for electrochemical finish machining. The power source (as a rule, pulse one) and the electrolyte (usually $NaNO_3$) differ from those for EDM.

The production of the razor shaving head at one of the divisions (DAP) of Royal Philips Electronics, where a production line that is completely equipped with ECM technology is employed [84], is a good example of electrochemical finish machining of shaped parts in mass production.

12.6.5 ECM in Biomedical Engineering

Application of ECM in the field of biomedical engineering is well known [64, 85–87]. Stainless steel plates for use in bone surgery and artificial joints (e.g. hip

joint) of special alloys, valves of extremely tough cobalt-based alloy used as a replacement in human hearts and a number of other metallic body inserts were produced using ECM. Owing to its good biocompatibility, titanium is widely used for dental and orthopedic implants and for biomedical microsystems. Titanium also has great potential for the fabrication of microfluidic networks and other microsystems such as micropumps for biomedical applications or for microchemical reactors. For these purposes, special methods of ECM of titanium were developed. The biomedical performances of titanium implants depend on their surface topography. C. Madore and D. Landolt used through-mask electrochemical micromachining (ECMM) to fabricate three-dimensional high-precision topographies on titanium [86]. A new four-step method for ECMM of titanium is presented in Ref. 87, which implies anodic oxidation of the titanium substrate, patterning of the oxide by excimer laser irradiation in air, electrochemical dissolution of the metal from the irradiated areas, and ultrasonic cleaning (to remove the protruding oxide film). The method is applied for fabricating 3D model structures, surface structuring of a cylinder and machining of a complex two-level architecture.

12.6.6 Subsea Application

Special equipment was worked out for ECM and combined processes in a subsea environment for acquiring metal construction defective zone samples for detailed studies in the laboratory or for the total removal of the defective zone (for instance, crack zone) to prevent the defect from growing, drilling small holes at the ends of the crack tip to prevent tip propagation and producing large diameter holes, and removal of entangled wire ropes [88–90]. The method and means may be used for welding repair in hard-accessible places of underwater constructions. Seawater may be used as an electrolyte; however, in some cases, it is advantageous to use concentrated $NaNO_3$ solution in order to accelerate the process of metal anodic dissolution. It is necessary to use the electrolyte flow systems.

12.6.7 Electrochemical Micromachining (ECMM)

In recent years, ECMM has developed intensively in accordance with the general tendency toward the development of various methods of micromachining for manufacturing small-size parts or structural elements with a characteristic dimension of about 1 mm.

In the first approach to the ECMM, the required localization is achieved by the use of inert photoresist masks on the WP areas, which should not be dissolved. This method has some specific features (for instance, the problem of mask undercutting and island formation) and was considered, for example, in Ref. 91.

In the second approach, the required localization of metal dissolution is achieved by other means. Some of them were considered in the previous sections. Here, TEs of several types are used [92–97]: (1) An electrode with the shape and dimensions corresponding to the required cavity in the WP; (2) a TE in the form of a wire or needle, which penetrates into the WP at the expense of its local anodic dissolution and moves according to a certain program, in order to obtain a given contour; and (3) a high-speed jet beam of electrolyte from a micronozzle.

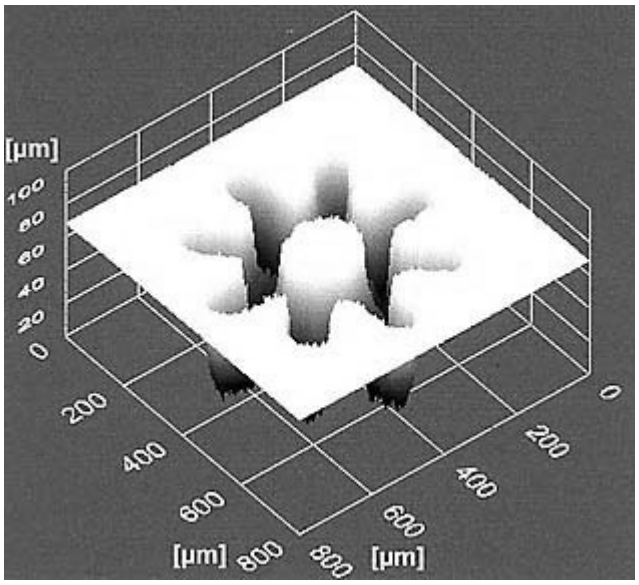


Fig. 24 A microgear pattern produced by micro-ECM on a chromium substrate using a small tip electrode [92].

Figure 24 gives a microgear pattern (700 µm in diameter, 70 µm in depth) produced by micro-ECM on a chromium substrate using a small tip electrode [92].

The removal of very small amounts of metal from the WP is an important feature of ECMM as well as other methods of micromachining, and the accuracy of machining should be characterized by micron and submicron dimensions. The first feature considerably simplifies the operating process, because high currents and high-amperage power sources become unneeded. The amount of reaction products formed in the interelectrode space is small; no high pressure or electrolyte overheating issues arise. All this enables one to significantly decrease interelectrode distances down to about one micron or even less, thus enhancing the machining accuracy. Acidic solutions are frequently used as electrolytes in order to prevent the formation of solid dissolution products in the very small interelectrode space. Mini equipment (such equipment was described, for example, in Refs 92, 93, 97) and a small amount of electrolyte can be used to realize ECMM. This extends the capabilities of ECM to more expensive electrolytes and simplifies the problem of environmental concerns of the machining processes.

In Sect. 12.2.1, the mechanism of enhancing the degree of dissolution localization using pulse ECM in solutions exhibiting an increasing dependence of the current efficiency on the current density was briefly described. In these solutions, the pulse-on time, which is the best from the viewpoint of conventional ECM, was not too short (tens and hundreds of microseconds), that is, the productivity of ECM was not decreased too strongly. Such rather long pulses appeared to be optimal provided that the charge Q^* consumed in the preliminary processes (charging of

electrical double layer, passivation, oxygen evolution, the initial stage of metal activation) prior to the onset of high-rate metal dissolution is large.

In the ECMM, it is necessary to remove very small amounts of metal; this overshadows the problem of providing a high productivity of the process, which is of great importance. In this case, the electrolytes in which the passivation of the dissolving metal is virtually absent (e.g. chloride solutions) can be used, and the methods of improving the accuracy on the basis of maintaining the WP areas, which should not have been machined, passive, are unsuitable. Application of such electrolytes is also important because for some metals, copper, for example, no sufficiently simple and convenient practical electrolyte composition, in which the condition $E^{\#} > E_{ox}$ is fulfilled, has been found.

After imposing the voltage on the system, a current is generated, which is consumed for charging the nonequilibrium electrical double layer and for the Faradaic process; in the absence of passivation, the latter primarily involves metal dissolution. In the very initial period of time, a major fraction of the current is consumed by the charging process; however, with time, the charging current decreases, and dissolution becomes the predominant process. Upon completion of charging, the entire current is consumed solely by the Faradaic process.

Under these conditions, in order to obtain a high degree of accuracy of electrochemical reproduction, the pulse parameters (pulse-on time, the amplitude) should be chosen such that, on the area with the minimum gap (the minimum voltage drop in the solution and, correspondingly, the maximum current), a sufficient charge will be consumed by metal dissolution, that is, by the machining. On the areas with larger gaps, the electrical charge in the pulse is smaller to the extent that only a small fraction is consumed by metal dissolution. Dissolution is localized on the areas with the minimum gap. Very short pulse-on times correspond to this mechanism of localization.

In recent studies [95, 96], the method of ECMM with supershort nanosecond pulses has been developed. This method enables one to perform machining with submicron resolution. It was shown by approximate calculations and by experiment that the machining conditions (primarily, short pulses) can be chosen, such that on the adjacent area points, which are offset by only several microns from the machining zone boundary, virtually no WP dissolution occurs, whereas in the machining zone (immediately facing the working surface of the TE) considerable dissolution proceeds. The authors of these works [95, 96] explain this fact by comparing the time constant τ of double-layer charging (the product of electrolyte resistance by the double-layer capacity) with the voltage pulse duration t_p. A dissolution rate, which is sufficient for certain practical use, is obtained at a required pulse amplitude provided τ does not considerably exceed t_p. The electrolyte resistance and, consequently, τ are proportional to the interelectrode distance; therefore, t_p can be chosen such that a sufficient degree of double-layer charging and the required metal dissolution rate will be achieved only at the minimum interelectrode distance (the minimum resistance), that is, in the working zone. Additionally, attention was drawn to the exponential dependence of electrochemical reaction rate on the polarization.

The rate of machining in these modes is hundreds times lower than under

conventional ECM conditions; however, a very high resolution allows its application to many important modern technologies, where it is necessary to remove small amounts of metal with great accuracy.

12.6.8 Hybrid Processes

The devising of hybrid methods is a line of development of nontraditional methods of material machining. In these methods, two types of action on the machining surface are simultaneously used. Together with electrochemical dissolution, most frequently, the abrasive (electrochemical grinding), the electroerosion (electrochemical discharge machining), the ultrasonic, or the laser action is used. The hybrid processes are applied with the aim of completely using the advantages of both types of action and minimizing the disadvantages of each method owing to their mutual effect. This may be achieved at the optimal ratio between the contributions of each mechanism to metal removal.

At a sufficiently high voltage, the electrolyte in the IEG boils and a local electric breakdown of the vapor-gas layer occurs (the electrochemical discharge machining). At the area where the breakdown occurs, a large amount of heat evolves and a portion of the melted metal is removed owing to the high pressure in the discharge zone. A cavity forms at this place. The next discharge occurs at another area. At each instant of time, one or a few electric discharges occur in the gap. The rest of the surface dissolves anodically. The machined surface is formed by the superposition of a great number of electroerosion cavities and by anodic dissolution assisting surface smoothing. The higher the amplitude and the duration of the voltage pulse, the higher the metal removal rate (the machining productivity); the larger the size of the electroerosion cavity, the poorer the surface finish and the stronger the TE wear. An optimal combination of the contributions of electrochemical and electroerosion mechanisms of metal removal allows one to combine high accuracy of shaping, which is characteristic of the EDM, with good surface finish, which is typical of ECM, at a high productivity of the process.

In electrochemical grinding, the mechanical removal of both the passive anodic film and the metal proceeds concurrently with the anodic dissolution. In this method, normally electrolytes, in which the metal dissolution is localized only on the areas of abrasive depassivation, are used. This enhances the machining accuracy in relation to the ECM. As compared with mechanical grinding, the combined method is characterized by a significantly lower tool wear and a high productivity.

Hybrid methods of ECM are detailed in reviews [98–103] and in many papers.

12.7 Surface Layer Integrity after ECM

ECM is performed without physical contact between the tool and the WP in contrast to mechanical cutting and without strong heating in the machining zone as in EDM. Therefore, no surface metal layer with mechanical and thermal distortion, compressive stresses, and cracks form in ECM. However, sometimes intergranular attack (superalloys) or hydrogen embrittlement (titanium alloys) of the surface layer occurs in ECM. This may degrade the performance of the machined parts. In order to eliminate the intergranular attack and hydrogen embrittlement, the electrolyte composition, temperature, and the electrical conditions must be optimized

[6, 104]. If this appears to be insufficient, special operation of mechanical hardening is performed after ECM, or a surface metal layer 0.05-mm thick is mechanically removed.

Surface finish after the ECM strongly depends on the electrolyte composition and the machining parameters (i, E_a, the temperature, and so forth). When a metal in the active state dissolves, etching normally takes place. After machining in the limiting or overlimiting current ranges, bright or even polished finishes are obtained. However, in this case, the risk of the formation of grooves on the surface in the direction of electrolyte flow (''striation'') is incurred. An increase in the temperature can lead to a decrease or an increase of the roughness, depending on the WP material. An increase in i, as a rule, reduces the roughness. The problem of surface roughness evolution in the course of metal anodic dissolution has been analyzed in a number of reviews [4, 105, 106].

12.8 Control of Electrolyte Composition and Environmental Problems

In the course of ECM, as a result of electrochemical reactions proceeding on electrodes and chemical reactions proceeding in the IEG, the electrolyte composition changes. This can impair the ECM performance and lead to environmental pollution. In many works, for example [107–117], these processes are studied and the methods of electrolyte regeneration and prevention of environmental pollution are developed.

In the interelectrode space, the dissolved metals, in most cases (Ni, Fe, Al, Ti), convert into insoluble hydroxides. To separate sludge from the liquid phase, several methods are used. The simplest method is sedimentation in a tank. However, this method is not always satisfactory with respect to productivity and quality of purification. To accelerate coagulation, special substances (polyacrylamide, for example) are sometimes added to the electrolyte. The schemes and equipment for flotation cleaning have been elaborated [114]. Various types of hydrocyclones, filtration devices, including automated setups of continuous action to service a number of machines, centrifuges, and various setups for combined methods of cleaning are used. If the electrolyte pH after the ECM is lower than 9, an alkali is added to the electrolyte for converting the compounds of most metals (nickel, for example) into precipitates. Upon sludge separation, the electrolyte composition is corrected, in particular, with respect to the pH value. The methods of sludge utilization have been described in the literature [115, 116]. In some cases, it is necessary to remove excess passivating nitrite ions, which can accumulate in the solution.

The removal of six-valent chromium from an electrolyte is a special problem. Its toxic compounds (e.g. CrO_4^{2-}) remain in the dissolved state, though a fraction of chromium is adsorbed by the sludge. To transform dissolved chromium into sludge, substances converting Cr(VI) into Cr(III) are added to the spent electrolyte. Na_2SO_3, $NaHSO_3$, and $FeSO_4$ are commonly used as the reductants. By introducing barium salts into the electrolyte, Cr(VI) can be directly precipitated in the form of insoluble $BaCrO_4$. A fraction of Cr(VI) is reduced in the electrolyte by the reaction with Fe^{2+} and NO_2^- yielded by the cathodic reaction (5).

The machines for ECM are supplied with ventilation for the removal of gaseous reaction products, primarily, hydrogen.

12.9 Conclusions

ECM has a significant place among the methods of machining of conducting materials. It is applied in cases where alternative methods (especially the traditional ones) encounter problems: for operations which are time and hand-labor consuming. ECM is feasible for the machining of difficult-to-cut materials (high-strength, tough, hard, and brittle alloys, superalloys), complex-shaped parts, and the machining in the difficult-to-reach areas. In some cases, ECM enables one to perform unique operations, for example, drilling of very small, deep, curved holes.

The ECM processes are in progress, spread into new areas of modern technologies, for instance, micromachining.

References

1. A. E. De Barr, D. A. Oliver, (Eds.), *Electrochemical Machining*, Macdonald, London, 1968.
2. V. A. Golovachev, B. I. Petrov, V. G. Filimoshin et al., *Elektrokhimicheskaya Razmernaya Obrabotka Detalei Slozhnoi Formy*, Mashinostroenie, Moscow, 1969.
3. J. F. Wilson, *Practice and Theory of Electrochemical Machining*, Wiley, New York, 1971.
4. J. A. McGeough, *Principles of Electrochemical Machining*, Chapman & Hall, London, 1974.
5. F. V. Sedykin, *Razmernaya Elektrokhimicheskaya Obrabotka Detalei Mashin*, Mashinostroenie, Moscow, 1976.
6. V. A. Shmanev, V. G. Filimoshin, A. Kh. Karimov et al., *Tehnologiya Elektrokhimicheskoi Obrabotki Detalei v Aviadvigatelestroenii*, Mashinostroenie, Moscow, 1986.
7. A. I. Dikusar, G. R. Engelhardt, V. I. Petrenko et al., *Elektrodnye Protsessy i Protsessy Perenosa v Elekrtokhimicheskoi Razmernoi Obrabotke*, Shiintsa, Kishinev, 1983.
8. E. Rumiantsev, A. Davydov, *Electrochemical Machining of Metals*, Mir, Moscow, 1989.
9. A. D. Davydov, J. Kozak, *Vysokoskorostnoe Elektrokhimicheskoe Formoobrazovanie*, Nauka, Moscow, 1990.
10. E. V. Tuck, in *Industrial Electrochemical Processes* (Ed.: A. T. Kuhn), Elsevier, Amsterdam, 1971, pp. 263–326.
11. A. H. Meleka, D. A. Glew, *Int. Met. Rev.* **1977**, *22*, 229–252.
12. J. Bannard, *J. Appl. Electrochem.* **1977**, *7*, 1–29.
13. J. P. Hoare, M. A. LaBoda, in *Techniques of Electrochemistry* (Eds.: E. Yeager, A. J. Salkind), Wiley, New York, 1978, Vol. 3, pp. 48–141.
14. J. P. Hoare, M. A. LaBoda, in *Comprehensive Treatise Electrochemistry* (Eds.: J. O. V. Bockris, B. E. Conway, E. Yeager et al.), Plenum Press, New York and London, 1981, Vol. 2, pp. 399–520.
15. K. P. Rajurkar, D. Zhu, J. A. McGeough et al., *Ann CIRP* **1999**, *48*, 567–579.
16. W. Guseff, British Patent 335003, 1929.
17. S. Cowper-Coles, *Electro-Chem. Metallurg.* **1903**, *3*, 203–204.
18. T. L. Lievestro, *Metal Handbook*, 9th ed., Machining ASM International, Metal Park, 1989, Vol. 16, pp. 533–555.
19. P. G. Pott, *Proceedings of International Symposium ElectroMachining (ISEM-9)*, The Japan Society of Electro-Machining Engineers, Nagoya, 1989, 146–150.
20. D. G. Risko, A. D. Davydov, *J. Am. Soc. Mech. Eng.* **1993**, *64*, 701–711.
21. M. Datta, *IBM J. Res. Dev.* **1993**, *37*, 207–226.
22. A. D. Davydov, *Russ. J. Electrochem.* **1991**, *27*, 835–847.
23. A. D. Davydov, V. M. Volgin, V. V. Lyubimov, *Russ. J. Electrochem.* **2004**, *40*, 1230–1265.
24. V. I. Petrenko, A. I. Dikusar, O. A. Arzhintar, *Elektron. Obrab. Mater.* **1976**, (4), 14–17.
25. Yu. N. Petrov, V. V. Parshutin, *Elektron. Obrab. Mater.* **1976**, (4), 22–25.
26. A. D. Davydov, E. N. Kiriyak, V. D. Kashcheev, *Sov. Electrochem.* **1978**, *14*, 352–355.
27. R. A. Mirzoev, N. K. Svechkopal, A. G. Lukina, *Fiz. Khim. Obrab. Mater.* **1971**, (5), 131–133.
28. P. Novak, B. Sajdl, I. Rousar, *Electrochim. Acta* **1984**, *30*, 43–49.
29. J. Kozak, K. Lubkowski, K. P. Rajurkar, *Trans. NAMRI/SME* **1997**, *25*, 159–164.

30. A. Zaytsev, I. Agafonov, N. Gimaev et al., *J. Mater. Process. Technol.* **2004**, *149*, 419–425.
31. A. V. Rybalko, S. I. Galanin, A. G. Atanasyants, *Elektron. Obrab. Mater.* **1993**, (3), 3–6.
32. A. Davydov, *1st International Conference on Machining and Measurements of Sculptured Surfaces*, Krakow, 1997, pp. 317–330.
33. V. S. Krylov, A. D. Davydov, J. Kozak, *Sov. Electrochem.* **1975**, *11*, 1075–1194.
34. A. L. Krylov, *Dokl Akad Nauk SSSR* **1968**, *178*, 321–323.
35. J. M. Fitz-Gerald, J. A. McGeough, *J. Inst. Math. Appl.* **1969**, *5*, 387–408.
36. R. Hunt, *Int. J. Number. Methods Eng.* **1990**, *29*(6), 1177–1192.
37. C. M. Elliott, *Free boundary problems: theory and applications*, Boston, 1983, pp. 505–511.
38. C. M. Elliott, *J. Inst. Math. Appl.* **1980**, *25*, 121–131.
39. A. A. Lacey, *IMA J. Appl. Math.* **1985**, *34*, 259–267.
40. H. Tipton, *V International Conference on Machine Tool and Research*, Birmingham, 1964, pp. 509–522.
41. J. Kozak, *Arch. Bud. Masz.* **1967**, *15*, 240–256.
42. A. Kh. Karimov, V. V. Klokov, E. I. Filatov, *Metody Rascheta Elektrokhimicheskogo Formoobrazovaniya*, Kazanskii Universitet, Kazan, 1990.
43. J. Deconinck, *Current Distributions and Electrode Shape Changes in Electrochemical Systems, Lecture Notes in Engineering, Vol. 75*, Springer-Verlag, Berlin, 1992, pp. 281.
44. G. A. Prentice, C. W. Tobias, *J. Electrochem. Soc.* **1982**, *129*, 72–78.
45. Y. Zhou, J. J. Derby, *Chem. Eng. Sci.* **1995**, *50/17*, 2679–2689.
46. H. Shirvani, A. Shirvani, I. Mirzaee, *II International Conference on Advances in Production Engineering. Part II*, Warsaw, 2001, pp. 343–352.
47. M. J. Noot, A. C. Telea, J. K. M. Jansen et al., *Comput. Vis. Sci.* **1998**, *1*, 105–111.
48. M. Purcar, J. Deconinck, L. Bortels, *J. Mater. Process. Technol.* **2004**, *149*, 472–478.
49. L. Bortels, J. Deconinck, P. Van Ransbeeck, *J. Mater. Process. Technol.* **2004**, *149*, 486–492.
50. C. S. Chang, L. W. Hourng, C. T. Chung, *J. Appl. Electrochem.* **1999**, *29*, 321–330.
51. C. S. Chang, L. W. Hourng, *J. Appl. Electrochem.* **2001**, *31*, 145–154.
52. V. M. Volgin, A. D. Davydov, *Russ. J. Electrochem.* **2001**, *37*, 1197–1205.
53. V. M. Volgin, O. V. Volgina, A. D. Davydov, *Russ. J. Electrochem.* **2002**, *38*, 1059–1067.
54. V. M. Volgin, O. V. Volgina, A. D. Davydov, *Comput. Biol. Chem.* **2003**, *27*, 185–196.
55. J. Kozak, A. D. Davydov, *Sov. Electrochem.* **1983**, *19*, 769–786.
56. G. A. Prentice, C. W. Tobias, *J. Electrochem. Soc.* **1982**, *129*, 78–85.
57. Z. H. Qiu, H. Power, *J. Appl. Electrochem.* **2000**, *30*, 575–584.
58. D. Adalsteinsson, J. A. Sethian, *J. Comput. Phys.* **1995**, *122*, 348–366.
59. V. M. Volgin, V. V. Lyubimov, *XII International Symposium for Electromachining (ISEM XII)*, Aachen, 1998, VDI BERICHTE, pp. 523–532.
60. V. M. Volgin, V. V. Lyubimov, *2nd International Conference on Advances in Production Engineering. Part II*, Warsaw, 2001, pp. 299–308.
61. V. K. Jain, K. P. Rajurkar, *Precis. Eng.* **1991**, *13*(2), 111–124.
62. J. Kozak, *J. Mater. Process. Technol.* **1998**, *76*, 170–175.
63. V. M. Volgin, V. V. Lyubimov, *J. Mater. Process. Technol.* **2001**, *109*, 314–319.
64. A. Goyer, S. Moehring, *Modern Machine Shop*, 1988, April, pp. 62–69.
65. A. N. Zaytsev, *Elektron. Obrab. Mater.* **2001**, (6), 71–79.
66. G. Bellows, J. B. Kohls, *American Machinist*, Special Report 743, 1982, March, 173–188.
67. M. S. Ahmed, A. Duffield, SME Technical Paper MS89-816, Society of Manufacturing Engineering, Dearborn, 1989.
68. G. F. Benedict, *Nontraditional Manufacturing Processes*, Marcel Dekker, New York, Basel, 1987.
69. J. W. Janssen, SME Technical Paper MS89-818, Society of Manufacturing Engineering, Dearborn, 1989.
70. *Prod. Eng.* **1967**, *38*(12), 126.
71. G. Bellows, *Tool Manuf. Eng.* **1967**, *58*(5), 32–34.
72. J. Kozak, K. P. Rajurkar, R. Balkrishna, *J. Manuf. Sci. Eng. Trans. ASME* **1996**, *118*, 490–498.
73. V. V. Zenin, V. P. Kondrat'ev, Yu. M. Vodyanov et al., *Elektron. Obrab. Mater.* **1975**, (5), 85–87.
74. M. Uchiyama, T. Shibazaki, *J. Mater. Process. Technol.* **2004**, *149*, 453–459.

75. E. J. Weller, (Ed.), *Nontraditional Machining Processes*, 2nd ed., Society of Manufacturing Engineers, Dearborn, 1984.
76. L. K. Gillespie, *Deburring Technology for Improved Manufacturing*, Society of Manufacturing Engineers, Dearborn, 1981.
77. H. Schein, SME Technical Paper MR87-154, Society of Manufacturing Engineering, Dearborn, 1987.
78. T. Wang, *Int. J. Electr. Mach.* **1996**, (1), 9–12.
79. D. Ball, in *Cost Guide for Automatic Finishing Processes* (Ed.: L. J. Rhoades), Society of Manufacturing Engineers, Dearborn, 1981.
80. S. Sakai, T. Masuzawa, S. Ito, *Proceedings of International Symposium Electro-Machining (ISEM-9)*, The Japan Society of Electro-Machining Engineers, Nagoya, 1989, pp. 155–158.
81. T. Masuzawa, S. Sakai, *Ann. CIRP* **1987**, *36*, 123–126.
82. F. Klocke, A. Karden, *XII International Symposium for Electromachining (ISEM XII)*, Aachen, 1998, VDI BERICHTE, pp. 375–385.
83. B. Lilly, J. Brevick, C.-S. Chen, *Trans. NAMRI/SME* **1997**, *25*, 153–158.
84. H. S. J. Altena, *J. Mater. Process. Technol.* **2004**, *149*, 18–21.
85. J. Bannard, *J. Appl. Electrochem.* **1975**, *5*, 89–90.
86. C. Madore, D. Landolt, *J. Micromech. Microeng.* **1997**, *7*, 270–275.
87. P.-F. Chauvy, P. Hoffmann, D. Landolt, *Appl. Surf. Sci.* **2003**, *208–209*, 165–170.
88. D. Clifton, M. B. Barker, R. W. Gusthart et al., *Underwater Technol.* **1991**, *17*(4), 1–9.
89. J. A. McGeough, *Ocean Ind.* **1983**, *18*, 282.
90. T. J. S. Green, *Underwater Syst. Des.* **1986**, *8*(5), 28–30.
91. M. Datta, D. Landolt, *Electrochim. Acta* **2000**, *45*, 2535–2558.
92. M. Suda, N. Watanabe, K. Furuta et al., *2nd International Workshop on Microfactories (IWMF 2000)*, Fribourg, 2000, pp. 163–166.
93. B. Bhattacharyya, S. Mitra, A. K. Boro, *Rob. Comput. Integr. Manuf.* **2002**, *18*, 283.
94. T. Masuzawa, T. Takawashi, *XII International Symposium for Electromachining (ISEM XII)*, Aachen, 1998, VDI BERICHTE, pp. 1–15.
95. R. Schuster, V. Kirchner, P. Allongue et al., *Science* **2000**, *289*, 98.
96. M. Kock, V. Kirchner, R. Schuster, *Electrochim. Acta* **2003**, *48*, 3213.
97. J. Kozak, K. P. Rajurkar, Y. Marrar, *J. Mater. Process. Technol.* **2004**, *149*, 426–431.
98. A. D. Davydov, *Russ. J. Electrochem.* **1994**, *30*, 871–881.
99. A. D. Davydov, R. A. Mirzoev, D. G. Risko et al., *Surf. Eng. Appl. Electrochem. (Elektron. Obrab. Mater.)* **1995**, (2), 6–18.
100. A. D. Davydov, J. Kozak, *Surf. Eng. Appl. Electrochem. (Elektron. Obrab. Mater.)* **1993**, (3), 1–16.
101. J. Kozak, K. P. Rajurkar, *Adv. Manuf. Sci. Technol.* **2000**, *24*(2), 25–50.
102. J. Kozak, K. P. Rajurkar, *Adv. Manuf. Sci. Technol.* **2000**, *24*(3), 5–24.
103. J. Kozak, K. P. Rajurkar, *Trans. NAMRI/SME* **2001**, *29*, 421–427.
104. L. Strode, M. B. Bassett, *Wear* **1986**, *109*, 171–180.
105. D. Landolt, *Electrochim. Acta* **1987**, *32*, 1–11.
106. Yu. D. Gamburg, A. D. Davydov, Yu. I. Kharkats, *Russ. J. Electrochem.* **1994**, *30*, 422–443.
107. W. Konig, J. Friedrich, *Galvanotechnik* **1988**, *79*, 1796–1802.
108. W. Konig, J. Friedrich, L. Liebau, *VDI-Z.* **1991**, *133*, 42–44.
109. H. Yamamoto, *Denki Kagaku oyobi Kogyo Butsuri Kagaku* **1979**, *47*, 373–376.
110. H. Simon, *Ind. Prod. Eng.* **1979**, (1), 19–26.
111. C. van Osenbruggen, C. de Regt, *Philips Tech. Rev.* **1985**, *42*, 22–32.
112. A. I. Dikusar, A. I. Molin, O. B. Anan'eva, *Electron. Obrab. Mater.* **1984**, (6), 5–9.
113. L. V. Saltanovskaya, A. I. Dikusar, *Electron. Obrab. Mater.* **1990**, (4), 46–49.
114. N. G. Kharlan, *Electron. Obrab. Mater.* **1974**, (4), 85–87.
115. L. P. Fominskii, E. V. Gorozhankin, M. L. Dantsig et al., *Elektrofizicheskie I Elektrokhimicheskie Metody Obrabotki* **1981**, (11), 1–3.
116. V. Kh. Postanogov, T. I. Evgen'eva, V. I. Anokhin et al., *Elektrofizicheskie I Elektrokhimicheskie Metody Obrabotki* **1982**, (9), 4–5.
117. M. A. Lubnin, I. V. Trifonov, V. Kh. Postanogov et al., *Electron. Obrab. Mater.* **1984**, (1), 64–66.

13
Environmental Electrochemistry

Barry MacDougall, C. Bock and M. Gattrell
National Research Council of Canada, Ottawa, Ontario, Canada

13.1 **Introduction** **857**

13.2 **The Electrochemical Approach in Solution** **858**

13.3 **Electrochemical Treatment of Organic Toxins in Solution** **860**

13.4 **Cathodic Electrochemical Dehalogination** **866**

13.5 **Electrochemical Treatment of Anionic Inorganic Toxins in Solution** . **869**

13.6 **Cathodic Removal of Toxic Metal Ions from Solution** **871**

13.7 **Advanced Oxidation Technologies Employing Hydroxyl Radicals** . . . **872**

13.8 **Ion Permeable Membranes** **874**

13.9 **Electrochemical Remediation of Contaminated Soil** **876**

13.10 **Sensor Technology in the Environmental Field** **877**

13.11 **Concluding Remarks** **878**
References 879

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

13.1 Introduction

During the past 30 or more years, there has been an increasing scientific and political emphasis on environmental issues related to air, soil, and water pollution in all parts of the world. Societies for the most part now realize that serious pollution problems can adversely affect both their economic and health-related quality of life. Indeed, there are many who feel that environmental issues are a prime problem for mankind as we enter the twenty-first century, and that serious action has to be taken if we want to avoid major crises in the years to come. To address these issues, a large number of environmental technologies have come forward to either handle the existing pollution problems or prevent them from occurring in the first place. These include chemical and biological processes as well as physical treatment processes. Electrochemical intervention and treatment form an important subset of the former, and has become increasingly important in the environmental field during the past decade or so. A number of excellent reviews and books have recently appeared on the subject, and a representative few are given here [1–9].

Electrochemistry itself has a very long history in the energy and manufacturing sectors of society, and therefore has the necessary infrastructure to make a major contribution to solving environmental problems. Some of the factors that make electrochemistry an attractive approach to solving, or preventing, environmental problems are as follows: (1) the operation is possible under ambient temperature and pressure; (2) switching the system on or off is easy; (3) electrochemical systems avoid the need to use reactive chemicals with the possibility of sludge production and its necessary disposal; (4) electrochemical systems can be small, modular, and portable, that is, ideal for remote location applications; (5) a large inorganic electrochemical industry exists (e.g. chloralkali and metal refining) with excellent cell design and engineering know-how; (6) the electrochemical approach can be highly selective for particular reactions since the electrode potential determines which reaction(s) will occur. Indeed, the strength of the electrochemical approach is that the electrode potential applied to the anode and cathode is all that is required for oxidation and reduction reactions to occur, that is, there is no need to add highly reactive (and sometimes dangerous) chemicals. In view of all of these advantages for the electrochemical approach to environmental management, it is in a strong position to make even more impact on this

Encyclopedia of Electrochemistry. Edited by A.J. Bard and M. Stratmann
Vol. 5 *Electrochemical Engineering*. Edited by Digby D. Macdonald and Patrik Schmuki
Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. ISBN: 978-3-527-30397-7

multibillion dollar technology area. The present chapter will discuss those areas where electrochemistry has already had a significant impact and those where there is still a lot of potential for the future.

In the environmental field, electrochemistry can be employed to do the following: (1) monitor pollutants; (2) clean up and recycle spent solutions and/or ionic species in solution; (3) remediate problems that already exist by, for example, anodically mineralizing organic toxins to CO_2 and H_2O or cathodically removing dissolved metals; (4) develop a cleaner, more efficient manufacturing process. While the last approach is obviously the ideal one since the problems with pollution are addressed before they occur, it cannot be universally applied to all systems and processes. As a consequence, there will be the need for some technical intervention or remediation at some point, during the process itself and/or afterwards, to clean up any arising pollution problems. Besides electrochemistry, one needs to also consider a long list of other potential treatment technologies, including biological, UV/ozone/peroxide, supercritical water oxidation, sonolysis, photochemistry (really related to electrochemistry), activated-carbon adsorption, chemical oxidation or treatment, and both ion-exchange and standard membrane separation [10]. Each approach has its advantages and disadvantages, and often a hybrid approach combining several of the above (possibly with electrochemical treatment) is the most advantageous one. While electrochemistry itself may not always provide the complete solution to an environmental problem, it can often play the essential role in a hybrid approach.

In the present chapter, the focus will be on the applications of electrochemistry to solving environmental problems and issues encountered in both aqueous and nonaqueous solutions, as well as in soils. The use of electrochemistry in air purification and electrochemical sensing of pollutants will also be mentioned. Some of the practical barriers to the implementation of electrochemistry in the environmental field will be discussed. An attempt will be made to address both the fundamental research and applied engineering sides of the issue, with insight into the all-important question of scale-up economics. This is a critical issue if the developing electrochemical technologies are going to migrate from the research laboratory into the field.

13.2 The Electrochemical Approach in Solution

Electrochemistry offers a versatile, flexible approach to solving environmental problems in solution, as illustrated in Fig. 1. Either the anode or cathode of the electrochemical cell can be used, and even a combination of both in the case where, for example, cathodic dechlorination generates a substance that is more easily oxidized electrochemically. This is illustrated in Fig. 2 for the model compound pentachlorophenol (PCP), a substance with a high oxidation potential because of the presence of the electron-withdrawing chloride groups [11, 12]. Dechlorination of the PCP to the more easily oxidized phenol can be cathodically achieved with high current efficiency (CE) in a nonaqueous solution and low CE in an aqueous one (see the following text). It should be noted that the whole process requires a considerable number of electrons, something that will influence the economics of the system. Besides the obvious direct, heterogeneous reaction at either electrode, it is also possible to generate oxidizing or

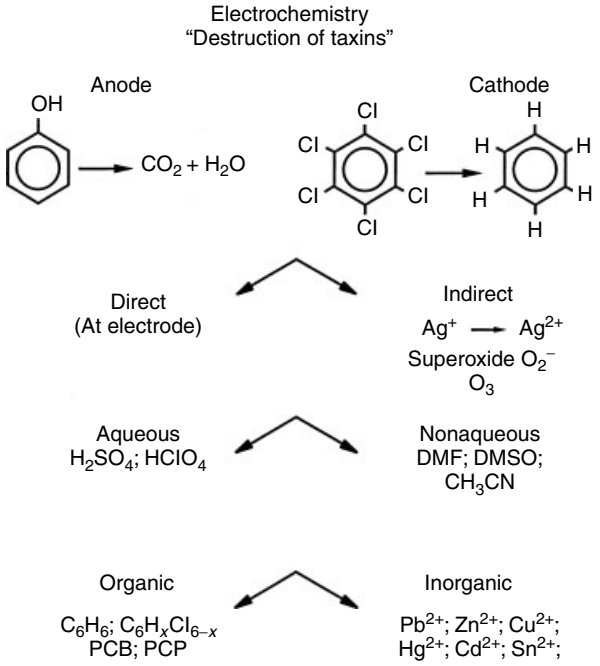


Fig. 1 Scheme illustrating the electrochemical destruction (or modification) of organic toxins at either the anode or the cathode and the significant scope of the electrochemical approach. For example, the electrochemical treatment can be achieved either directly at the electrodes or indirectly by involving solution species generated at the electrode(s); it can be applied to either aqueous or nonaqueous systems and the species to be treated can be either organic or inorganic toxins.

reducing species at the electrode, which can migrate into solution, giving rise to a homogeneous charge transfer. The mediating charge-transfer species can usually be reused over and over again, serving as a long-lasting electron-transfer shuttle vehicle. Examples include inorganics such as cerium and silver, and organics such as acetophenone and biphenyl. The direct and indirect electrochemical approaches can be employed in either aqueous or nonaqueous solutions, provided of course there is sufficient conductivity. The advantage of working in nonaqueous solutions is the much larger available potential range of operation, which can be more than 5.0 V and is usually limited by the breakdown potential of the electrolyte (compare the usual 1.23 V range in aqueous systems). As an example, the complete dechlorination of hexachlorobenzene to benzene can only be achieved in nonaqueous solutions because of the very high cathodic potentials needed to remove the chlorides from the ring, especially the last few. A complication sometimes encountered in nonaqueous solutions is the potential need for a source of protons (to complete the dechlorination),

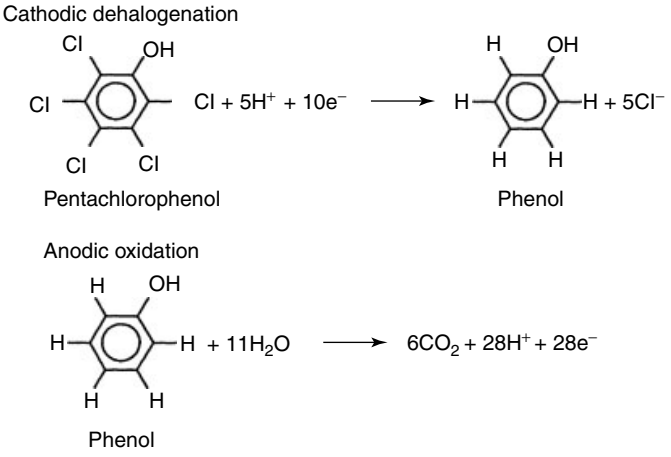


Fig. 2 Electrochemical detoxification/destruction reactions for treatment of pentachlorophenol where an initial cathodic dechlorination is followed by anodic oxidation of the more readily treated phenol.

which is the need for at least some water in the system. For this reason, it may be advantageous to have a small amount of water in the solution. The last part of Fig. 1 simply indicates that electrochemistry can be applied to the treatment of both organic and inorganic species, the organics being treated at either the anode or the cathode while the inorganics are mainly handled at the cathode. The purpose of Fig. 1 is to illustrate the wide range of the electrochemical approaches in treatment, separation, and/or recovery of pollutants in solution. It should also be mentioned that electrochemistry can be applied to both concentrated and very dilute solutions (the latter being something of a challenge); more will be said about this later.

13.3 Electrochemical Treatment of Organic Toxins in Solution

A considerable amount of attention will now be paid to the anodic destruction/modification of problem toxins, both organic or inorganic, in solution since this is a topic of great academic interest with considerable industrial potential at the present time. It also illustrates very well the important elements common to all electrochemical reactor systems, that is, nature of solution; geometry, activity, and lifetime of anode catalyst; optimum current and/or potential for system operation; design of electrochemical cell and overall treatment process (e.g. batch versus continuous, divided versus undivided cell); need for auxiliary equipment such as pumps and valves; and the need for nonelectrochemical pre- or posttreatment such as filtration and biotreatment, if a hybrid system is required [13, 14]. The overall cost of the system, broken down into capital and operating cost, will determine how competitive the process is, and whether the electrochemical treatment route has a chance of being chosen by industry to do the particular environmental

prevention or remediation job. As a general rule of thumb, the most expensive component of the electrochemical organic treatment system is the anode catalyst, whose cost is generally in the region of US$1200/m^2 for the DSA™ material (RuO_2/Ti). If a material like graphite could be used, the cost would fall greatly since it is ca. 10 times less expensive. The problem is that inexpensive anode materials will also have short lifetimes and therefore have to be replaced frequently. It is not uncommon for the active anode material to represent 60% or more of the entire treatment system [13, 14], and this means that it is important to investigate less expensive anode materials or to discover how to utilize existing materials much more efficiently so that the number of (meters)^2 required in the system decreases. For this reason, research into anode electrocatalysis for organic treatment systems has become extremely important and a great many papers have been published and patents issued, in the field recently.

The publications dealing with anodic destruction of organics at metal-oxide electrodes can effectively be divided into two main groups: (1) those of Johnson and others [15–23], who concentrate on the modified PbO_2 electrode system; (2) those of Comninellis and others [24–30], who deal with the modified SnO_2 system. The reason for using high-oxygen overvoltage anodes is to drive the electrochemical reaction into a potential range where oxidation of the organics is possible (during galvanostatic anodic treatment); correspondingly, the extent of parallel oxygen evolution decreases during constant potential treatment. In the extensive PbO_2 work of Johnson et al. [15–21], the central idea is that incorporation of an alter-valent ion such as Bi(5+) into the oxide lattice greatly facilitates oxygen transfer to the solution species (R) to be treated. [When reading these papers, one should be aware that there were three procedures for doping the films with Bi: (1) anodic adsorption as Bi(5+) ad atoms onto β-PbO_2 before the oxygen-transfer experiment; (2) having Bi (3+) in the solution during the oxygen-transfer experiment at the β-PbO_2 electrode; and (3) formation of the anode coating by deposition onto substrates such as gold from Bi(3+)/Pb(2+) solutions of varying ratios prior to the oxygen-transfer experiment. The latter approach was the more standard one and probably the most relevant so far as eventual scale-up is concerned.] The important reactions at play are

$$S[\,] + H_2O \longrightarrow S[OH] + H^+ + e^- \quad (1)$$

$$S[OH] + R \longrightarrow S[\,] + RO + H^+ + e^- \quad (2)$$

$$S[OH] + H_2O \longrightarrow S[\,] + O_2 + 3H^+ + 3e^- \quad (3)$$

(1) being discharge of water onto the oxide surface S, (2) being oxygen transfer to a reactant R in solution, and (3) being parallel oxygen evolution. The process is referred to as ''electrochemical incineration'' and ideally results in either complete mineralization of the organic, R, to CO_2 or its conversion to a less-toxic form. Short-term experiments were conducted on the oxidation of organics such as phenol, benzene, toluene, and *p*-benzoquinone, as well as inorganics such as cyanide, Cr(3+), and dimethylsulfoxide (DMSO) (as a model for the oxygen-transfer process). The influence of the presence of various dopants in the β-PbO_2, as well as the doping level, was investigated. The general observation was that the presence

of incorporated Bi(5+) had a beneficial influence on the oxidation reactions studied, so long as the amount incorporated did not exceed ca. 50% (a level at which the word ''doping'' should probably not be used). It is interesting to note that a similar beneficial effect was achieved by incorporating other ions such as Fe(3+), Al(3+), Cl^-, F^- and even acetate anions into the β-PbO_2 [16, 18], that is, it is not specific to Bi(5+), although much of the research was done on this system. It indeed appears that the dopant can be tailored to the reactant to be treated for ''best'' performance, for example, Fe(3+) in the PbO_2 best suits the oxidation of cyanide. The overall results were interpreted in terms of the incorporated alter-valent ion species generating surface defects in the β-PbO_2, which in turn generate active oxygen-transfer species such as the hydroxyl radical (OH*) (see Fig. 3). What is needed for a continuous oxidation reaction is the rapid generation of lots of these ''active-oxygen'' species, which are weakly bound to the surface and always ready to oxidize targeted solution species. The altered electrodes are also less prone to the well-known radical-ion-induced polymerization of phenolic compounds on the surface, meaning that they remain ''active'' for continuous oxygen transfer. In terms of the exact reaction mechanism, Johnson et al. have suggested that the Bi(5+) may provide adsorption sites for the organic [23], the maximum enhancement being near the 50/50 Bi:Pb ratio. In view of the fact that even incorporated species such as Cl^- and acetate can enhance the organic oxidation reaction at PbO_2 [16, 18], the importance of organic adsorption at the dopant sites is open to question and further research is needed in this area. Nevertheless, the notion that doping an oxide such as β-PbO_2 with ions such as Bi(5+) can significantly increase the rate of oxidation of dissolved solution species has important environmental implications for remediation of waste and other waters. It should be noted, however, that most of the experimental work discussed earlier was short term in nature, with data mainly obtained using the technique of cyclic voltammetry. As a consequence, there is little knowledge of the long-term performance characteristics of these electrodes under the demanding conditions that would be encountered in the field. Long-term stability of the electrodes is obviously extremely important from an economic as well as an environmental acceptance point of view, since the possibility

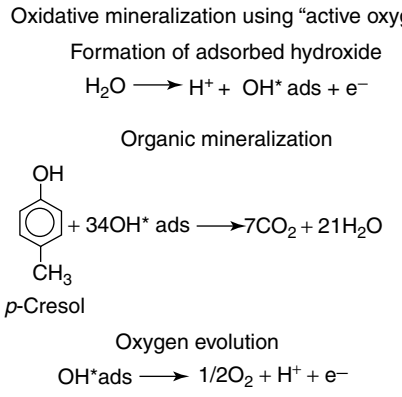


Fig. 3 Reaction scheme for the oxidative mineralization of an organic toxin using an ''active-oxygen intermediate'' that is involved in the oxygen-evolution reaction.

of leaching Bi and/or Pb into treated waters is not attractive.

Along the same lines, a considerable amount of research has been done over the past 20 years on another high oxygen-evolution overpotential oxide, namely, SnO_2 [24–30]. The standard dopant in this case is antimony, Sb(5+), [31–35] at a level of several atom %, the upper level being ca. 4% as dictated by its solubility in the SnO_2. The oxide coatings are generally formed on a Ti substrate by the spray-pyrolysis technique, with prior conditioning of the Ti and/or formation of interlayers (such as IrO_2) needed for good coating adhesion [36, 37]. A number of researchers have worked on this particular system, notably Comninellis [24–27] and others from Switzerland [28, 29]. They noted that Sb-doped SnO_2 is highly efficient at oxidizing organics such as phenol [24] and 1,4-benzoquinone [25] when compared with Pt and IrO_2. As was the case with Bi-doped PbO_2, the oxidation mechanism has been suggested to involve a generated hydroxide species on the oxide surface, which facilitates mineralization of the solution organics [30]. The presence of the Sb(V) provides the SnO_2 with electrical conductivity and also appears to assist in the generation of an active OH* species (see Fig. 3). The trick appears to be in forming an adsorbed (physisorbed) hydroxyl species on the oxide surface, that is, $MO_x + H_2O \rightarrow MO_x(\cdot OH) + H^+ + e^-$, which does not interact with the existing oxide to increase its oxidation state, that is, $MO_x(\cdot OH) \rightarrow MO_{x+1} + H^+ + e^-$. It has been suggested that it is the OH* species, which is capable of mineralizing organics to CO_2, and not the higher oxide that is believed to be involved with incomplete oxidation of the organic [30]. It should be noted, however, that a partial oxidation of an organic to a less-toxic form would itself be environmentally beneficial and indeed require fewer electrons. It should also be noted that recent studies using small organic molecules (e.g. oxalic or formic acid) has shown that adsorption of the organic can be importantly involved in the oxidation process [38, 39]. Furthermore, it is known that OH* radicals are not needed for complete and efficient oxidation of oxalic acid to CO_2. Indeed, good results have been obtained for anode materials exhibiting relatively low overpotentials for oxygen evolution. This means that even if OH* radicals are involved in some organic oxidation processes, this involvement is not universal.

In light of the observations by Comninellis and others, it is quite possible that one important role of the Sb-dopant in the SnO_2 lattice is to hinder the conversion of $MO_x(\cdot OH)$ to MO_{x+1} by acting as a ''third-body'' or impurity in the SnO_2 lattice. The same may well be true for Bi (as well as the other dopants) in PbO_2. By acting as an impurity in the oxide lattice, the dopants can prevent, or at least slow down, the stoichiometric incorporation of new oxygen into the lattice to generate a higher and/or more stable oxide. It may also be that the dopants generate defects in the oxide, which are more likely to favor the formation of active radical OH* species. Whatever the reason, the positive effects of such dopants on organic (and even inorganic) oxidation is very real. As was the case with the β-PbO_2(Bi), the service life of the Sb-doped SnO_2 electrodes under real operating conditions [37], as well as the cost of their manufacture and production reliability, must be known. Indeed, one has to be able to consistently produce large amounts of the coated-electrode material with a controlled chemical composition and morphology. While this is presently

true for materials such as the DSA Cl_2 (RuO_2/Ti) and DSA O_2 (IrO_2/Ti), it is not as yet the case for many of these emerging new anode catalyst formulations. This situation could very well change if and when a large market for these materials is established.

A relatively new anode material for organic destruction that has received considerable attention in the last few years is B-doped diamond [40–46], a highly corrosion-resistant material that has a very wide potential window for water stability and is capable of complete mineralization of solution organics with high current efficiencies. The mechanism is believed to be similar to that encountered with SnO_2(Sb) and PbO_2(Bi) electrodes [44], but without the associated anode stability problems. The major problem with the B-doped diamond has been cost; however, this has been decreasing in the past several years and there are some indications [47] that it may now be approaching about two times that of standard DSA (which itself is approximately US$1200/$m^2$). Also, very long-term testing under ''real'' operating conditions still needs to be performed. Another fairly recent candidate anode material is Permelec, manufactured by Permelec Electrode Ltd., Kanagawa, Japan. It is actually a coating of β-PbO_2 deposited on a thinner layer of α-PbO_2, with additional interlayers separating the coating from the Ti substrate. Special fibers or powders in the outer β-PbO_2 layer reduce stress in the coating and make it very durable (see Ref. 48, the patent). These electrodes have been shown to be as catalytically active as standard PbO_2 electrodes but with stabilities that are 100 to 1000 time higher; they are also very stable in organic electrolytes [49]. At the present time, the cost of Permelec is quite high; however, this could change if the market dictates large-scale production.

As an example of the dramatic influence of various anode catalyst materials on system performance, the results for electrochemical treatment of solutions from a printed-circuit board (PCB) copper plating bath are shown in Fig. 4. In its freshly prepared state, the bath consisted of 2 M H_2SO_4, Cu^{2+}, and standard organic brighteners, levelers, and carriers, which gave a total organic carbon (TOC) content of about 400 ppm. During long-term plating operation in the field, the TOC increases to ca. 2000 ppm because of dissolution of photoresists and masks, and so on. At 2000 ppm, the quality of the deposited Cu deteriorates considerably and it is necessary to bring the TOC level down again. Adsorption of the solution organics onto active carbon (in a side-holding tank) followed by environmentally unfriendly land filling is one possible treatment route. However, electrochemistry offers an alternative whereby the solution can be treated in an ''in situ'' manner to bring the TOC levels back to an acceptable range. The results in Fig. 4 were obtained with an AB micro cell from ElectroCell, Sweden, employing a Nafion cation-exchange membrane as separator [50]. It is immediately obvious that the selection of anode material is critical to the efficiency of organic removal from solution, the high-oxygen overvoltage PbO_2 being by far superior to either the Pt or IrO_2 (the latter is the DSA oxygen electrode). In longer-term tests, the PbO_2 (Fig. 4), experiences some corrosion problems that may be overcome by using the special PbO_2 coating, Permelec, discussed earlier. The plating-bath system described here has two obvious advantages over standard waste streams in terms of ease of treatment, that is, the high conductivity of the solution and the rather high

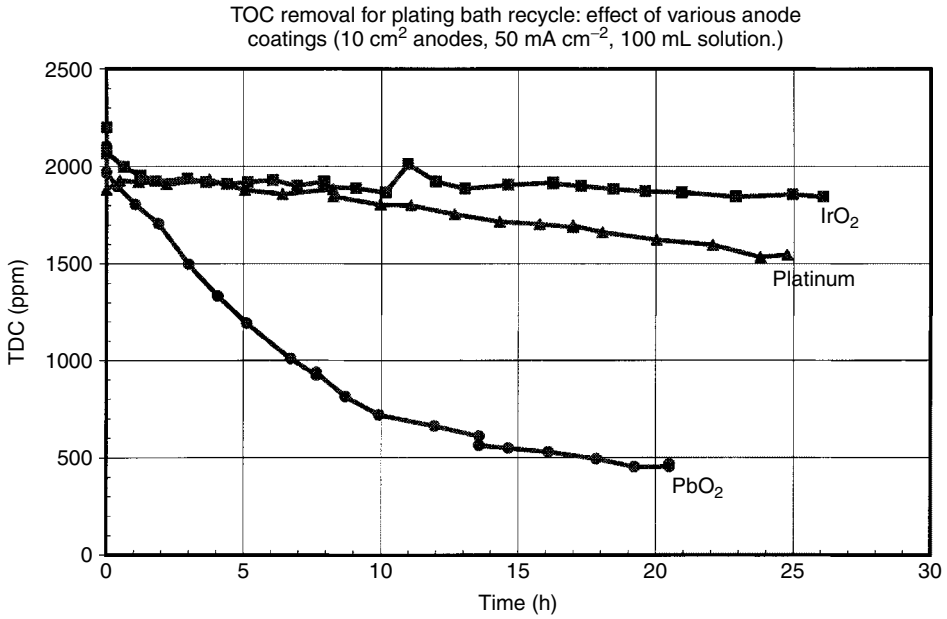


Fig. 4 Influence of time of anodic galvanostatic treatment of a 2 M H_2SO_4 copper plating bath solution containing approximately 2000 ppm of organics (see text). Note that 10 h of charging corresponds to the passage of 180 000 C L^{-1} of solution. Also note that it is impossible to establish the current efficiency (CE) for the organic removal since the exact identity of the organics in this complex solution are unknown.

concentration of the organics present (a disadvantage is the very aggressive nature of the 2 M H_2SO_4). In the treatment of standard waste streams and contaminated groundwater, large volumes of solution need to be treated and it is not really economical to add supporting electrolyte to increase solution conductivity. One way around this problem is to preconcentrate the dilute toxin beforehand and possibly transfer it to a more conductive solution. More will be said about this approach in Sect. 12.4.

A novel cell design has recently been suggested for the treatment of organics and microorganisms in wastewater. It is known as the Zappi™ system, and has been demonstrated to treat a wide range of organic pollutants in various waste streams [50]. While only a limited amount of information has as yet been published on this new technology, which is in the process of being patented, it has been learned that the cell is of an inexpensive, simple design with no gaskets. It consists of an open stack of electrodes between two end plates, with recirculating fluid flow; the fluid "leaks" through the sides of the system and is caught in a tray at the bottom for recirculation. It appears to work better in the presence of added NaCl (possibly due to indirect oxidation by generated hypochlorite), but this is apparently not necessary for its successful operation. A fairly large-scale facility is being presently demonstrated in South Korea (at a landfill leachate site) and results seem to be promising. More information on this new system should be forthcoming in the near future. Other commercial

(or noncommercial) ''anodic'' wastewater treatment systems are based on the indirect, homogeneous approach of electrochemically generating strong oxidants. One such system, sometimes referred to as the ''silver bullet'', depends on the anodic generation of Ag(2+) from a nitric acid solution of Ag(+) [51, 52]. Reactive oxygen species such as *OH radical (in solution) are the most probable ''active'' species, for example, $Ag(2+) + H_2O = Ag(+) + H^+ + {}^*OH$; the Ag(2+) is regenerated again at the anode. Because of the cost of silver, as well as the fact that it precipitates in the presence of halide ions in solution, other redox systems such as Fe(2+)/(3+) and Co(2+)/(3+) have also been investigated [53, 54]. There has been considerable success in laboratory testing of these mediated treatment technologies, and a few companies worldwide are marketing systems that employ the technology (e.g. see Ref. 6). It should also be mentioned that standard oxidants such as ozone, hydrogen peroxide, and chlorine dioxide, which can all be generated electrochemically, are obviously of great interest in the field of wastewater treatment and are outlined in Chapter 7 of Ref. 6.

13.4 Cathodic Electrochemical Dehalogination

Another viable electrochemical method for treating halogenated organics is cathodic dehalogenation, whether directly at a cathode surface or indirectly using charge-transfer mediators generated at the cathode that migrate into the solution. The use of a somewhat novel and simple approach, based on zero-valent iron, will also be mentioned in the following text. The general overall reaction can be represented as

$$RX + H^+ + 2e^- \longrightarrow RH + X^- \quad (4)$$

and involves the generation of an organic anion radical, R^-, which undergoes protonation; the halide ion is generally chloride. While there has been some research on the cathodic dechlorination approach in aqueous solutions [55], the CE tends to be very low (<1%) because of parallel hydrogen evolution. Research conducted by Schmal et al. on the dechlorination of PCP in aqueous solutions at graphite/carbon cathodes [55–57] showed the reaction going through tetra-, tri-, di-, monochloro- and all the way to phenol, that is, conversion to a less toxic and more biotreatment-compatible product. While a US patent was issued for the process in 1984 [58], the method does not appear to have been widely used in the field. Another approach for removal of chlorinated phenols from aqueous waste streams involves their electrochemically induced condensation onto high surface area carbon anodes using only one electron per toxin molecule [59, 60]. Under appropriate conditions of applied current, solution pH, and toxin concentration, the removal from solution onto the anode occurs with a CE of 100% and toxin levels can be decreased from >100 ppm to <1 ppm. One of the major advantages of using this anodic gettering approach is that the loading of the carbon electrodes is substantially enhanced over that observed with simple physical adsorption; for example, the enhancement factor is 91 in the case of PCP, 208 for 4-cresols, and 219 for phenol [59]. These very substantial ''adsorption'' enhancements mean that the carbon adsorbent can be used for much longer times in service before it is saturated and needs to be replaced. It should be remembered, however, that this particular approach simply transfers the toxins from solution to

adsorbent, that is, it must still be treated at some future time. Possible posttreatments include (1) standard landfilling of the carbon (not ideal); (2) transfer of the carbon absorbent to a nonaqueous solution for cathodic dechlorination (see the following text); or (3) electrochemical and/or chemical desorption of the toxins into a smaller volume of solution where the ''more concentrated'' toxins can be treated more easily than was the case with the very dilute solutions (a type of preconcentration step). The latter is possible because the toxins can be desorbed from the carbon electrode by application of a cathodic current (as shown by C^{14} tracer experiments [59, 60]). A patent has been recently issued on this treatment technology [61], and potential applications in the field are being investigated. As with many of these emerging technologies, problems arise with ''real'' waste solutions (as opposed to synthetic laboratory ones), which can contain large amounts of suspended matter as well as a wide range of organic and inorganic pollutants. Pretreatments are often necessary, and the treatment process must be able to handle a feed that can change dramatically from day to day, or even hour to hour.

Cathodic dechlorination can be carried out effectively in nonaqueous solvents such as acetonitrile (AN), DMSO, or dimethylformamide (DMF), the available potential range being much extended over that for aqueous systems. As an example of direct dechlorination, Fig. 5 shows the cathodic sweep behavior for a solution of 5 mM hexachlorobenzene in AN at a sweep rate of 50 mV s^{-1} [62]. It is obvious that the dechlorination proceeds in a stepwise fashion from stage 1 (conversion to pentachlorobenzene) to stage 6 (conversion to benzene). The final stage of the reaction (i.e. 6) is somewhat masked because of the breakdown of the electrolyte at the high cathodic potentials encountered. There are definite advantages in performing cathodic dechlorinations by an alternative, indirect route, using solution organic mediators such as anthracene, acetophenone, phenanthridine, and so on. The mediator, A, is reduced at the cathode to give an anion radical, A^-; one or more of these radicals migrate into solution and

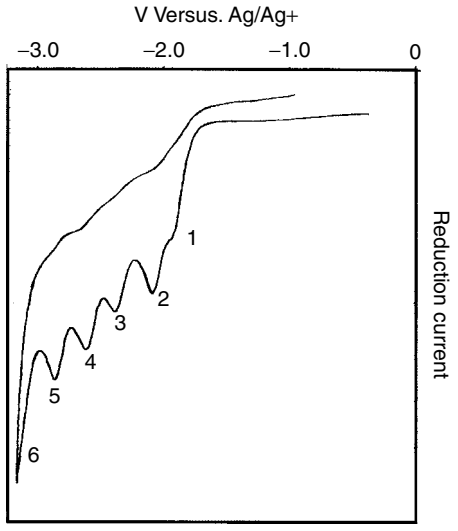


Fig. 5 Cyclic voltammogram of hexachlorobenzene at a glassy carbon electrode in acetonitrile containing 0.1 M tetraethylammonium tetrafluoroborate (TEABF4). Sweep rate is 50 mV s^{-1}. The cathodic peaks represent the stepwise dechlorination process.

interact with the organochloride, thereby reducing the latter and regenerating the mediator molecule, which is then ready for further reductions [63]. The mediated reduction occurs at a lower cathodic potential than for direct reduction, that is, the reaction is catalyzed by the mediator. Since the dechlorination reaction occurs homogeneously in solution at a somewhat less cathodic potential, possible filming of the cathode is avoided and the reaction can proceed for long periods of time at the same potential (under galvanostatic conditions) with a high CE. Proper selection of the mediator and the reaction conditions can ensure either complete or partial dechlorination (as desired) with little or no loss of the added mediator. One potentially important application of this approach is in the dechlorination of polychlorinated biphenyls (PCBs) arising from, for example, spent electrical transformers. It has been clearly demonstrated that with the use of biphenyl as mediator, mixtures of PCBs can be completely dechlorinated with a high CE to yield biphenyl and some reduced products from biphenyl [62]. Using biphenyl as a mediator has the advantage that it is also one of the products of the dechlorination reaction, thus ensuring a steady supply of mediator and making later product separation simpler. A real operating system would have to use something like a dissolving metal anode for the counter electrode reaction (e.g. Mg), and this complicates the cell design because of precipitation of insoluble salts of the metal anode. Nevertheless, in areas where the electricity costs are not excessively high, this electrochemical method for treating PCBs holds promise.

As an alternative to working in nonaqueous solutions like DMSO and AN, considerable research has been conducted on the subject of electrodes modified by surfactant films as well as microemulsions in aqueous solutions [64]. Both topics have been reviewed in some detail by J.F. Rusling [65, 66], with extensive references therein. The idea is to be able to work under essentially nonaqueous conditions in an aqueous environment by coating the working electrode with an insoluble surfactant film, with or without incorporated metal catalysts, or forming microemulsions in solution that shield the ''reactant-to-be treated'' from the water environment. The microemulsion can then interact with the electrode surface, and charge transfer can occur to the reactant to be treated. PCBs and other organohalides have been dechlorinated in microemulsions [67–69], and organohalides have been preconcentrated and catalytically reduced at surfactant-coated electrodes [70–73]. Work has also been conducted (especially by T. Franklin et al.) on the electrooxidation of a variety of organics including organosulfur and organohalide compounds in aqueous sufactant systems using oxidation catalysts such as barium peroxide, copper (2+) oxide, and bismuth (3+) oxide [74–76]. The higher oxidation state of the oxide drives the organic oxidation reaction and is regenerated in the lower oxidation state, ready for reuse. Until now, these approaches have mainly produced very good fundamental research articles; however, there is promise that the techniques can be used in real aqueous system to achieve toxin treatments that were previously considered impossible in systems based on water.

An interesting and novel approach to the dechlorination of organics in soil and groundwater involves the use of ''zero-valent iron''. While mainly designed for treatment of soils, it will be discussed here because of the close relationship

to cathodic dechlorinaton. Iron metal, in the form of filings or granules, is simply buried in the ground after excavating some of the native soil. The iron undergoes oxidation to Fe^{2+}, that is, the standard corrosion reaction, with the counter cathodic reaction being dechlorination of certain organochlorides:

$$RCl + H^+ + 2e \longrightarrow RH + Cl^- \quad (5)$$

The iron is in the form of a permeable reactive wall that is installed in the ground and through which the contaminated groundwater must flow. Considerable research has been done in this field by Gillham and others [77–80] and its applicability in the field has been demonstrated [81, 82]. The technique can be used to treat a wide range of aliphatic organohalides (e.g. CCl_4, C_2Cl_6, CH_2Cl_4, vinylchloride, etc.), the Fe serving as a true catalyst in lowering the overvoltage for dehalogenation. Pollutant levels have been decreased from 100s of ppm to ~1 ppm level. The approach does not, however, appear to work for halogenated aromatics. The technique has been proven on both pilot demonstration and full-scale installations in the United States and Europe. Investigations are now being made on its effectiveness at depths >40 feet, the limit of standard excavation methods, using special engineering injection methods. The attraction of the zero-valent iron dechlorination method is that it is extremely simple, requiring no outside electrochemical control of the reactions, and it would appear to have great future potential. For further information on the practical applications of the technology, the reader is referred to the following representative website: www.beak.com/eti.html.

13.5 Electrochemical Treatment of Anionic Inorganic Toxins in Solution

Anodic oxidation has also been used to modify and/or destroy toxic inorganic species in solution, the most common and important example being cyanide from electroplating, surface finishing, and gold-mining operations. This topic has taken on added importance recently in light of a number of high-profile cyanide spills from large holding ponds (or lakes) that have contaminated both surrounding soils and rivers. Point source cyanide emission limits are usually set in the 1 ppm range, as opposed to the hundreds of parts per million usually encountered in the holding ponds. The electrochemical treatment approach, while competing with chemical oxidation methods such as chlorination, ozonation, SO_2/air (Inco Process) [83,] and hydrogen peroxide treatment (see e.g. Ref. 84), has the advantage of not producing large amounts of sludge [85]. The general reaction mechanism has been known for a very long time and involves conversion to the more benign cyanate, that is,

$$CN^- + 2OH^- \longrightarrow CNO^- + H_2O + 2e \quad (6)$$

the actual mechanism probably involving hydroxyl radicals, as was the case with the anodic oxidation of organics. Indeed, Johnson has published a number of papers dealing with the anodic oxidation of CN^- at Fe^{3+}-doped PbO_2 electrodes [86, 87]. As was the case for organics, doping of the oxide electrodes enhances the rate of cyanide ion oxidation. Besides reaction (6), other possible anodic reactions are

$$4OH^- \longrightarrow O_2 + 2H_2O + 4e^- \quad (7)$$

$$CNO^- + 4OH^- \longrightarrow CO_3^{2-} + 0.5N_2 + 2H_2O + 3e^- \quad (8)$$

The CE for the cyanide destruction process is obviously determined by the extent of the parallel reaction (7), a factor which becomes much more important as the cyanide levels drops to below 100 ppm [88]. Complete mineralization of the CN^- by reaction (8) removes all traces of the offending species, but at the expense of considerably more charge consumption. It may be reasonable to stop the reaction at the CNO^- stage, something which is quite feasible since reaction (8) occurs with more difficulty than (6) [89] at a potential of almost 0.5 V more anodic [87]. At a pH $\leq$ 10, the cyanate ion hydrolyzes on its own according to the following reaction:

$$CNO^- + 2H_2O \longrightarrow NH_4^+ + CO_3^{2-} \quad (9)$$

In general, waste cyanide solutions contain metals such as copper, and possibly also thiocyanate (SCN^-), making for some very interesting ''solution-complex'' chemistry. For example, the copper can exist in such forms as $Cu(CN)_3^{2-}$ and $Cu(CN)_2^-$, along with the free cyanide (CN^-) and SCN^-, the latter also being capable of interacting with solution species [90]. If these solutions are electrochemically treated under the correct conditions, the metal solution component, for example, Cu, can be collected at the cathodic (see Sect. 6), while the anode is being used to destroy the CN^- and SCN^- compounds. (This again highlights the theoretical strength of the electrochemical approach where both anodic and cathodic processes can be simultaneously used in a beneficial way.) The solution chemistry of such systems is obviously very complicated, and great care is needed to avoid a swing to lower pH values where the liberation of HCN gas is possible. Many side reactions, as well as precipitation and polymerization reactions involving the CN^-, are possible and can significantly influence the conditions required for optimal operation of the electrochemical reactor. Along with the direct electrochemical route for treating cyanide solutions, there is also the indirect method as illustrated by the halide-mediated electrooxidation [91, 92]. In this case, chloride is added to the electrolyte-generating hypochlorite, OCl^-, at the anode. The hypochlorite oxidizes the CN^- and CNS^- to CNO^- in a series of fairly complex reactions, with regeneration of the Cl^- for use in the next cycle. While this approach has some attractive features, one needs to consider the serious material corrosion problems in the presence of added Cl^-, as well as the possible liberation of Cl_2 gas during operation.

In the case of the direct electrochemical approach, while the electrolysis conditions are less severe, the selection of the appropriate electrode material is still very important, and further reading on the use of stainless steel [93], platinum [94], graphite [95], doped SnO_2 [92], doped PbO_2 [86, 87, 96], and so on, is suggested. The economic viability of the electrochemical treatment approach is influenced in no small way by the cost and lifetime of the anode material; this can easily make or break the field implementation of the process. Some authors have used high-surface area, porous anodes for cyanide treatment in order to combat the problems of mass-transport limitations so evident at cyanide concentrations below 100 ppm [88]. That system consists of a reticulated vitreous carbon porous anode that was activated for cyanide oxidation by the deposition of some copper oxide. The process looks very promising at the laboratory scale,

with the cyanide level decreasing from 100 ppm to <1 ppm with excellent current efficiencies. The result clearly illustrates the importance of activating the anode material for the desired reactions as well as the importance of electrode geometry – in this case the use of a three-dimensional anode structure.

13.6 Cathodic Removal of Toxic Metal Ions from Solution

Cathodic electrochemistry has been used for many decades in the electrowinning and electrorefining of metals, as well as in the electroplating and electroforming fields. Indeed, this is one of the largest industrial areas for inorganic electrochemistry (along with the production of chlorine/caustic for the pulp and paper industry). It is therefore an attractive approach for the treatment and recovery of both heavy and transition metals dissolved in waste streams, not only to recover the value of the dissolved metals but also to make the streams acceptable for safe discharge into the environment or reuse in a manufacturing process [97]. The solutions to be treated may arise from plating and photographic operations, etching and other chemical processes, mining operations, and so on. The more standard approach of chemical precipitation of the metals by the addition of a base simply shifts the problem to that of having to deal with the sludge, which itself has to be treated. It is important to remember that there is both an economic (recovered value of metals often in their pure form as well as regeneration of, e.g., a spent etchant) and environmental benefit to the electrochemical treatment approach, and that the dollar value of the former is obviously much easier to calculate than that of the latter. Streams to be treated contain a wide range of dissolved metal concentrations, from <1 ppm to several percent by weight. The challenge is to economically treat the low-concentration solutions where the cathodic deposition process becomes limited by mass transport of the metal ion to the electrode, that is, there is a limiting current for the process. The two standard ways of increasing the limiting current are by (1) increasing the surface area of the electrode; and (2) increasing the turbulence in the solution so as to effectively decrease the diffusion-layer thickness. To achieve the former, one can use a three-dimensional porous electrode design, for example, in the form of foam or felt, as opposed to the standard planar, flat-plate design, or a packed (or fluidized) bed electrode system [98]. Examples of the use of this approach include the Chemelec, Retec, and enViro cells, [99], each having their own advantages and disadvantages in terms of ease of operation and ability to treat down to very low metal ion concentrations. With continuous operation, the three-dimensional (3-D) electrode material will eventually become loaded with deposited metal, which must be recovered in a separate in situ or ex situ operation [98, 100]. This step is necessary for both recovery of the deposited metal (possibly in a more concentrated solution) and regeneration of the three-dimensional cathode, which could eventually become plugged by the deposit. Of the three cell systems mentioned above, the enViro Cell is most compatible with low metal ion concentrations (10–50 ppm) and can bring the level down to <1 ppm in a single-pass operation.

The rotating cylinder electrode (RCE) system handles the problem of low

limiting current values by rotating the cathode (usually in the form of a stainless-steel cylinder) at high speeds [101, 102]. While this technology does not have the industrial maturity of the three-dimensional systems mentioned earlier, it does have one very attractive feature, that is, that the deposited metal can be, from time to time or continuously, scraped off using a stationary wiper without interrupting the process. The dislodged metal deposit can then be removed from the bottom of the cell for processing to a useful form. The morphology of the deposited metal is important for both its ease of removal and contribution to local turbulence, the ideal one being loosely adherent and rough as opposed to being uniform and compact. This is the reverse of the situation with the three-dimensional cell where a loose deposit would more quickly plug the porous structure and require replacement or regeneration. As with all electrodeposition processes, the nature of the deposit will depend on the applied current (or potential), the concentration of metallic ions in solution, and their state of complexation (which in turn depends on the anions present in solution). It should also be noted that the chemical makeup of the waste streams may indeed be quite variable with both time and place; hence the electrochemical and process conditions needed to effectively remove the metallic components may not stay constant. It is also important to remember that real, industrial solutions are normally more complex than model, synthetic ones prepared for testing in the laboratory; in the end, experimentation with real solutions is essential. For more information on the fundamentals and applications of the RCE system, the reader should consult Refs. 103–106 (and the references therein).

Another treatment approach that should be considered consists of preconcentrating the metal ions by passing the waste stream through an ion-exchange column, followed by elution into a lower volume, more concentrated solution [107]. At the higher metal ion concentrations, the limitations discussed earlier are not so severe and a wider range of cell designs (even the simple and relatively inexpensive plate and frame one) can be employed [108]. This reference gives examples of the combined ion-exchange/electrowinning technology applied to the treatment and removal of copper, zinc, and nickel from large-scale industrial process waste streams. Another approach is to use ion-exchange membranes in an electrochemical cell system, that is, to use the electrodialysis approach, to remove metal ions from a waste stream and concentrate them for further treatment [109]. While the process is best known for desalination both to recover NaCl from salt water and to produce potable water, it has also been used to treat nickel from an electroplating system [110]. A number of companies are involved in such preconcentrating approaches, including Electrochemical Design Associates, Inc. (Berkeley, CA), Prosep (Pickering, Ont), U.S. Filter, and so on.

13.7 Advanced Oxidation Technologies Employing Hydroxyl Radicals

In this section, we will be briefly dealing with remediation technologies commonly referred to as ''advanced oxidation technologies'' (AOTs) [as processes (AOPs)]. The common factor in all of these technologies appears to be the generation of

very reactive $OH^{\cdot}$ radicals, and there is thus similarity to the direct and indirect electrochemical destruction treatments discussed in Sect. 12.10. While the AOTs are not for the most part electrochemical technologies, they will be mentioned in this chapter because of their importance to pollution treatment and their implications for electrochemical processes (both in a competitive and complimentary vein). In Sect. 12.3, the electrochemical anodic destruction of pollutants was discussed in terms of the generation of very reactive $OH^{\cdot}$ radicals at or near the electrode surface along with the possible adsorption (or physisorption) of the pollutant. Two of the best-known generators of hydroxyl radicals in solution are ozone, O_3, and hydrogen peroxide, H_2O_2. Each can be generated electrochemically – H_2O_2 at a cathode and O_3 at an anode, or simply added to the solution in liquid (H_2O_2) or gaseous (O_3) form. It is usually necessary to assist in their decomposition to ''OH radicals'', by, for example, using ferrous ions with H_2O_2 in the case of the Fenton's Reagent:

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH^- + {}^{\cdot}OH \quad (10)$$

Another route to $^{\cdot}OH$ generation is ''homogeneous photolysis'' using UV light to irradiate the H_2O_2 and/or O_3 in solution. It is obviously necessary for the solution to be able to transmit the UV radiation, that is, it should not be too opaque and/or have suspended matter. Besides $^{\cdot}OH$, many other radical and reactive species are possible from the $UV/H_2O_2O/O_3$ system, and considerable research has been conducted in identifying the various species and their reactions with organic species in solution [111, 112]. The homogeneous treatment technology has been successfully applied over several decades now, and a fairly large number of companies have been established to exploit and market the technology worldwide [113].

In the alternative heterogeneous photolysis approach, a semiconductor catalyst (usually based on TiO_2) is employed in a high-surface area form. The activation process depends on the generation of electron/hole pairs at the semiconductor/solution interface when the irradiation energy exceeds the band gap energy of the semiconductor. The irradiation energy source can either be direct solar radiation or that from a UV-lamp source. The photogenerated holes give rise to $^{\cdot}OH$ radicals while the counter-generated electrons can be used to reduce inorganic pollutants [114] or simply drive standard cathodic reactions such as H_2 evolution or O_2 reduction. The principles upon which the approach is based are those of the well-established science of photoelectrochemistry, in this case applied to an advanced oxidation process for environmental remediation in the dry or wet state (see e.g. Ref. 115). As with the homogeneous approach, a wide variety of organic pollutants have been successfully treated by heterogeneous catalysis, and further reading in this major subject area is recommended [112, 113, 116]. In the gas phase, success has been reported in the use of photosensitive TiO_2 coatings on, for example, ceramic bathroom tiles and auto windshields, for both air disinfection and maintaining clean surfaces [115, 117–119]. In the treatment of volatile organic pollutants, the heterogeneous method appears to be considerably more effective for treatment of pollutants in the gas phase in comparison to the aqueous phase. For this reason, there is considerable commercial interest in initially stripping the VOCs from water or soil and afterwards degrading them in the gas phase. While there

are examples of success with the aqueous-phase TiO_2 approach for both organics and inorganics such as H_2S and CN^- [120–124], the technology is not as yet mature (in comparison with homogeneous photolysis), and more work is needed at the engineering scale-up phase in terms of both economics and performance. Other AOTs for pollution treatment such as electrohydraulic cavitation/sonolysis and supercritical water oxidation are outside the scope of this chapter, and further reading on this topic is recommended [125].

As with all other treatment technologies, considerable attention has been paid to the pollutant degradation reactions and the extent to which it proceeds to complete mineralization. This is extremely important, especially when one considers the possibility of generating intermediates that may be more toxic than the starting material. Consequently, care has been taken to identify intermediates and do proper mass-balance calculations for the reaction. Hybrid approaches should also be considered, for example, combining the AOTs discussed earlier with more standard treatments such as GAC adsorption in a pre- or poststage. In the end, economic considerations will determine the approach that is used. The reader is referred to Chapter 8 of Ref. 112 for a detailed evaluation of the economics of the various treatments discussed earlier.

13.8 Ion Permeable Membranes

Ion permeable (or ion exchange) membranes play a major role in electrochemical synthesis and processing and have revolutionized electrochemical engineering during the past 30 or so years. They are often composed of an organic polymer onto which ion-exchange groups are grafted to provide the ion transport; the most common example is the dupont Nafion cation-exchange membrane. When placed between the anode and cathode compartment of an electrolysis cell, the membrane acts as a ''permeable'' separator between the anolyte and catholyte, and selectively dictates the transport processes between the two solutions [126, 127]. Membranes can be positioned individually between electrodes, or in a multimembrane configuration (called a *stack*) with alternating cation and anion exchange capabilities. The former is represented by the huge production of chlorine/caustic using Nafion membrane technology while the latter is represented by the classic electrodialysis systems for desalinating seawater and/or production of table salt. The attraction of using the electrochemical approach with multimembrane stacks is that only one set of anode and cathode electrodes is needed in the system, and up to 100 or more ion-selective membranes can be placed between them. In the environmental field, electrochemically driven membrane processes are used for the effective separation and concentration of unwanted species in solution, as well as the development of zero (or very low level) effluent discharge systems in the process chemical industry [128, 129]. As with all system components, the main factors affecting their use are cost, reliability, and lifetime.

Membranes, selective to either cation or anion transport, are employed in many of the electrochemical treatment systems already discussed in this chapter. For example, in the cathodic recovery of metals from waste streams, a cation-permeable membrane can prevent the migration of anions such as Cl^- from the catholyte to anolyte, where oxidation to Cl_2 can occur with

undesirable consequences. Another pertinent example is for the anodic treatment of organics in an acid copper plating solution (discussed in Sect. 12.3). It is obviously necessary to separate the anode and cathode compartments in order to avoid losing the solution copper via plating at the cathode. Use of an anion exchange membrane would permit cross-over, and therefore loss of the sulfate from the plating solution. The answer is to use a cation-exchange membrane as separator. Although a small amount (~2%) of copper will cross over (in addition to the protons), this can eventually be recovered from the cathode.

The electrochemically driven membrane approach can also be used in regenerating metal-contaminated acid from, for example, spent pickling baths. Various cell configurations can be employed, depending on the arrangement of the ion-exchange membranes; the metallic impurities may be collected at the cathode while (more importantly) the spent acid is collected in a ''regenerated'' state in one of the cell compartments [130]. This process is certainly environmentally superior to the more standard neutralization and disposal treatment. A similar electrochemical-membrane approach is used in the so-called ''salt-splitting'' process, whereby a salt such as Na_2SO_4 is converted to H_2SO_4 in one cell compartment and NaOH in the other [131]. The fact that a low value salt such as Na_2SO_4 can be split into its acid and base counterparts shows the strength of the technology, that is, caustic can be produced in a process that does not generate chlorine [132]. Both two- and three-compartment cells can, and have, been used in salt splitting; Fig. 6 shows the situation with a three-compartment cell.

A relatively new membrane technology uses the so-called ''bipolar membrane'' for water splitting [133–135]. It consists of an anion- and cation-exchange membrane glued together by a special intermediate layer, which appears to facilitate the water splitting reaction. They are used in many different configurations between parallel-plate anodes and cathodes, often with added monovalent ion-exchange membranes. Their power rests with the fact that the water splitting occurs at a lower voltage than is needed to generate protons and hydroxyls at anodes and cathodes, respectively. Also, the water splitting reaction can occur at very high rates in a bipolar membrane, the reaction possibly catalyzed by the intermediate binding layer, which has added metal catalysts. In a similar fashion to electrodialysis, numerous bipolar membranes can be stacked between two electrodes, and the systems can be used for recovery of acids from pickling baths, salt splitting, and so on. While there is great potential for the use of the bipolar membrane technology, it has not as yet made major inroads in the field (with the exception of a few isolated examples). This is because these membranes are expensive and are prone to the standard membrane problems such as fouling and degradation in harsh environments; this especially is true for the anion-exchange component of the bipolar and the binding glue in the center of the structure. In addition, the bipolar membranes are susceptible to deterioration if multivalent cationic species such as calcium, magnesium, and iron are present in the feed stream. These need to be removed to levels of <1 ppm by a pretreatment procedure and this obviously adds to the overall cost of the process system. At the moment, the industrial future of bipolar membrane technology is not as clear as that of the standard monopolar ion-exchange technology, in spite of their great promise and potential.

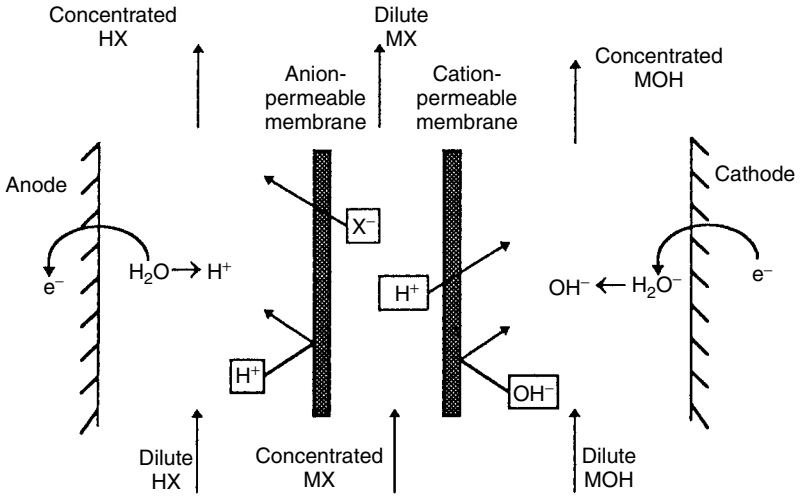


Fig. 6 Sketch of a three-compartment electrochemical cell, employing two monovalent ion-exchange membranes, used to "split" the concentrated salt solution MX (where M can be sodium and X can be sulfate, for example). Operating under optimum conditions, the cell arrangement will generate pure concentrated base in the cathode compartment and the corresponding acid in the anode compartment. (The illustration is taken from p. 14 of Ref. 124).

To complete this section, mention will be made of electrochemical ion exchange, referred to as EIX, systems. In such systems, the electric field produced between an anode and cathode is used to enhance an ion-exchange material's properties for the removal of selective ions according to their anionic or cationic nature [136–138]. An interesting application of the EIX principle is in the electrodeionization (EDI) of pure water to give ultrapure water for use in the semiconductor industry [139–141]. The system consists of alternating diluting and concentrating chambers bounded by anion- and cation-selective membranes; the diluting chamber is filled with ion-exchange resin beads much like the situation found in conventional mixed-bed ion-exchange systems. Application of the electric field facilitates the deionization of the water being treated, and at the same time helps in the continuous regeneration of the ion-exchange resins. This avoids the need for the standard resin regeneration using strong chemical additions, that is, a definite environmental benefit. A number of such systems are currently available where, in conjunction with reverse osmosis, very high purity (>18 Meg.) water is produced for a variety of applications.

13.9 Electrochemical Remediation of Contaminated Soil

Electrochemistry offers a simple, in situ approach for the treatment of both inorganic and organic contaminants in soils [142–144]. It is often referred to as "electrokinetic remediation", and is based on the placement and polarization of appropriate anodes and cathodes in the soil. The electrolyte is either the existing groundwater or an externally supplied fluid using

boreholes. The electric field in the wet soil moves anions to the anode and cations to the cathode; associated electroosmatic flow results in the movement and concentration of water and nonionic compounds toward the cathode (a process that has been used for decades to dewater soils and sludges [143]). While the role of electrochemistry in these systems is fairly simple, a good knowledge of geology and soil hydrology is important for the successful application of the treatment approach. Indeed, the transport processes are occurring through fine pores and capillaries in the treated soil, whose physical and chemical properties will depend on the nature and location of the soil. Electrochemical reactions occurring at the immersed anodes and cathodes produce acid and base regions, respectively, which move toward the opposite electrodes. Care must be taken in dealing with the moving base region since it can cause the precipitation of metal cations before they can be removed at the cathode [142]. This can be handled through neutralization by addition of an environmentally friendly acid such as CH_3COOH at the cathode region via a purging solution.

Besides being used for in situ remediation, the electrokinetic effect can also be used to contain, or hold, a plume of contaminants in the soil, which might otherwise spread to other areas. In this case, the applied electric field is adjusted so that the plume of contaminant is ''steady'', or immobile, until proper cleanup treatment can be arranged. This could be applied to soil contaminated with heavy organics, as found in fuel oil. In essence, the electrodes form a containment ring around the spill site, and the treatment is sometimes referred to as ''electrokinetic ring fencing'' (EKRF) [145].

The electrokinetic remediation approach has seen successful large-scale demonstrations and field tests during the past 10 years in North America and Europe [146, 147]. It has been shown to be applicable to removal of a wide range of soil toxins including Cr, Pb, Hg, Ni, Cl^-, CN^-, phenol, benzene, toluene, and so on. [148]. As a specific example, Monsanto's ''Lasagna'' process for soil remediation has received considerable attention during the past decade [149]. In this approach, the applied electric field moves the soil contaminants toward a series of vertical or horizontal treatment zones made, for example, of activated carbon. It has been successfully demonstrated for trichloroethylene (TCE), the process consisting of a ''first-phase'' adsorption on the active carbon followed by a ''second-phase'' in situ treatment to degrade the TCE using zero-valent iron (see Sect. 12.4) [150]. The process appears to be very attractive for low-permeability soils and clays where standard treatments (consisting of excavation, removal, and then treatment) are very expensive. It has been suggested that the Lasagna process could lower the standard treatment costs by an order of magnitude or more. More recently, Monsanto has changed its company focus and has decided to license this technology as opposed to exploiting it itself [151]. Other companies actively involved in electrochemical soil remediation include Geokinetics BV in the Netherlands, Electrochemical Design Associates, Inc. (EDA) in Berkeley, CA, and Electrokinetics, Inc. in Baton Rouge, LA.

13.10 Sensor Technology in the Environmental Field

In the environmental field, it is obviously important to determine the identity of

pollutants as well as monitor their concentrations in both the gaseous and liquid phases. In the scientific laboratory, this is accomplished by using analytical instruments such as HPLC and GCMS, in conjunction with TOC and COD measurements, to follow the nature and level of the pollutants as a function of treatment conditions. The crucial importance of such analysis hinges on the fact that, while complete mineralization of pollutants may not always be necessary, one certainly has to generate a product that is substantially less toxic than the starting material. Alternatively, there may be considerable benefit in electrochemically forming a product that can be easily handled by another treatment route, for example, the conversion of a nonbiodegradable pollutant into one that is readily biotreated [152]. It may, indeed, be economically advantageous to only do a partial electrochemical treatment in light of the number of electrons required for some reactions. This would then be followed by a subsequent alternative treatment (chemical, biological, etc.). To make these decisions with any degree of intelligence, it is obviously necessary to have information about the mechanism of the treatment reaction, that is, to perform product analysis and proper mass balances during the course of the reaction.

In field applications, the monitoring of the type and level of pollutants is also of great importance, and information is often required in real time rather than the several days of delay needed for sending samples to the laboratory. In response to this need, major advances have been made in the development and application of small-scale, portable sensors for in situ monitoring of pollutants [153–158]. Sensors can be of the solid-state, optical, bio-, and electrochemical types, examples of the latter being pH- and ion-selective electrodes. Their principles and modes of operation have been discussed in detail elsewhere [159], and the level of research in this scientific area has increased exponentially over the last few years. The most important factors governing the use of a sensor device are its cost, lifetime, reliability, specificity for the substance being measured, reproducibility, and so on. Sensors are becoming an integral part of pollution prevention and environmental remediation, and need to be considered as an integral part of any practical treatment system.

13.11 Concluding Remarks

In this chapter, we have attempted to give a broad ''overview'' of the environmental electrochemistry field. Recent approaches such as zero-valent iron dechlorination and AOTs, as well as long-established approaches such as cathodic removal of metallic ions, are discussed. The large range of treatments covered means that, out of necessity, a detailed description and discussion is at a minimum and the reader is encouraged to consult the extensive list of references given. But whether the particular electrochemical treatment technology is old or new, we hope that its actual and/or potential benefit to the field of environmental science and technology has been demonstrated. The science of electrochemistry would appear to be ideally suited to addressing environmental problems because of its versatility and ease of operation at ambient temperature and pressure. Electrochemistry offers the potential means to detect, identify, quantify, and treat pollutants, and, ideally, to even prevent their generation in the first place. The latter may

be achieved by either using electrochemistry as a cleaner synthesis and manufacturing route (if this is indeed feasible) or employing electrochemistry as part of an overall, hybrid production system in order to achieve cleaner manufacturing. In this chapter, most of the discussion has concerned the more standard electrochemical approach to detecting and treating pollution; nevertheless, electrochemistry can have a real role in preventing pollution in the first place.

As with all other environmental treatment and prevention technologies, the key to the utilization of electrochemistry is cost. It matters little how scientifically marvelous the technology is; if the cost is prohibitive, it will not be used. Factors influencing the cost of electrochemical technology include reaction CE, the electrocatalyst required and its lifetime, design, and construction of electrochemical reactor (plus auxiliary equipment such as pumps and valves), safety aspects of the process, necessity for pre- or post-treatment, and cost of electricity. In some instances, the latter may be the least concern, especially where electricity costs are in the range of 3.5 ¢/kwh. The electrode material and reactor system costs are critical since they represent a large fraction of the overall system cost. If a useful product can be obtained from the treatment method (e.g. metals from cathodic reduction), then this economic benefit can help to offset the treatment cost. The environmental benefit, while being most difficult to quantify than the economic return, is probably the most important factor for society as a whole.

In this chapter, a number of electrochemical treatment technologies have been discussed to address various environmental problems. Each problem will have its own specific circumstances that will decide which, if any, of these electrochemical technologies is appropriate from both a fundamental scientific and engineering economic point of view. It is necessary to know and understand the problem, as well as find potential solutions. Electrochemistry can play a very important role in pollution prevention and remediation, either on its own or as part of a bigger system, provided it is used with understanding and knowledge. Just as electrochemistry has had a major impact in the areas of corrosion prevention, mass production of bleaching agents, production, and refining of metals such as Al and Cu, and so on, the same is certainly possible in the environmental field.

References

1. G. Kreysa, K. Juttner in *Electrochemical Engineering and Energy* (Eds.: F. Lapicque, A. Storck, A. A. Wragg), Plenum Press, New York, 1995, pp. 225–270.
2. D. Simonsson, *Chem. Soc. Rev.* **1997**, *26*, 181–189.
3. K. Rajeshwar, J. G. Ibanez, G. M. Swain, *J. Appl. Electrochem.* **1994**, *24*, 1077–1091.
4. B. Fleet, *Bull. Electrochem.* **1993**, *9*, 497–501.
5. J. M. Fenton, *Interface*, Electrochemical Society, Pennington, 1994, pp. 38–44, Vol. 3, No. 1.
6. K. Rajeshwar, J. Ibanez, *Environmental Electrochemistry*, Academic Press, New York, 1997.
7. D. Genders, N. Weinberg (Eds.), *Electrochemistry for a Cleaner Environment*, The Electrosynthesis Company, East Amherst, 1992
8. C. A. C. Sequeira (Ed.), *Environmental Oriented Electrochemistry*, Elsevier, Amsterdam, 1994
9. P. M. Bersier, L. Carlsson, J. Bersier, *Top. Curr. Chem.* **1994**, *40*, 112–229.
10. See, eg., Abstracts from 1994 and 1995 *International Conference on Advanced Oxidation Technologies for Water and Air Remediation*, also special issues of J. Advanced Oxidation Technologies, London; 1999

11. S. I. Bailey, I. M. Ritchie, F. R. Hewgill, *J. Chem. Soc., Perkin Trans. 2* **1983**, *5*, 645–652.
12. S. D. Roso, M. Finkelstein, E. J. Rudd, *Anodic Oxidation*, Academic Press, New York, 1975, p. 273.
13. F. Goodridge, K. Scott, *Electrochemical Process Engineering*, Plenum, New York, 1995, Chapter 6.
14. D. Danly, D. Genders, D. Pletcher, Principles of electrochemical engineering and scale-up, *Short Course at '9th International Forum on Electrosynthesis in the Chemical Industry'*, Clearwater Beach, 1995
15. I. H. Yeo, D. C. Johnson, *J. Electrochem. Soc.* **1987**, *1341*, 1973–1977.
16. Y. L. Hsiao, D. C. Johnson, *J. Electrochem. Soc.* **1989**, *136*, 3704–3710.
17. K. T. Kawagoe, D. C. Johnson, *J. Electrochem. Soc.* **1994**, *131*, 3405–3409.
18. Y. L. Hsiao, J. E. Vitt, D. C. Johnson, *J. Electrochem. Soc.* **1992**, *139*, 377–380.
19. D. C. Johnson, N. Popovic, J. Feng et al., *The Electrochem. Soc. Proc. Vol.* **1995**, *95–26*, 176–188.
20. D. C. Johnson, L. L. Houk, J. Feng et al., *Environ. Sci. Technol.* **1999**, *33*, 2638–2643.
21. D. C. Johnson, H. Chang, J. Feng et al. in *Electrochemistry for a Cleaner Environment* (Eds.: T. L. Rose, E. Rudd, O. Murphy et al.), The Electrochemical Society, 1994, pp. 75–86, PV94–19.
22. N. B. Tahar, A. Savall, *J. Appl. Electrochem.* **1999**, *29*, 277–283.
23. N. D. Popovic, J. A. Cox, D. C. Johnson, *J. Electroanal. Chem.* **1998**, *456*, 202–209.
24. C. Comninellis, C. Pulgarin, *J. Appl. Electrochem.* **1993**, *28*, 108–112; **1991**, *21*, 703–708.
25. C. Pulgarin, N. Alder, P. Peringer et al., *Water Res.* **1994**, *28*, 887–894.
26. C. Comninellis, *Electrochimi. Acta* **1994**, *39*, 1857–1862.
27. C. Comninellis in *Water Purification by Photocatalytic, Photoelectrochemical and Electrochemical Processes* (Eds.: T. L. Rose, E. Rudd, O. Murphy et al.), The Electrochemical Society, 1994, pp. 75–86, PV94–19.
28. S. Stucki, R. Kotz, B. Carcer et al., *J. Appl. Electrochem.* **1991**, *21*, 99–104.
29. R. Kotz, S. Stucki, B. Carcer, *J. Appl. Electrochem.* **1991**, *21*, 14–20.
30. A. Savall, *Chimia* **1995**, *49*, 23–26.
31. C. Terrier, J. P. Chatelon, R. Berjoan et al., *J. Sol.-Gel. Sci. Technol.* **1997**, *10*, 75–82.
32. H. Cachet, F. Zenia, M. Froment, *J. Electrochem. Soc.* **1999**, *146*, 977–1013.
33. A. Messad, J. Bruneaux, H. Cachet et al., *J. Mater. Sci.* **1994**, *29*, 5095–5103.
34. H. Cachet, M. Froment, F. Zenia, *J. Electrochem. Soc.* **1996**, *143*, 442–448.
35. J. Bruneaux, H. Cachet, M. Froment et al., *Electrochim. Acta* **1994**, *39*, 1251–1258.
36. L. Lipp, D. Pletcher, *Electrochimi. Acta* **1997**, *42*, 10911111.
37. B. Correa-Lozano, C. Comninellis, A. DeBattisti, *J. Appl. Electrochem.* **1997**, *27*, 970–974.
38. C. Bock, B. MacDougall, *Electrochimi. Acta* **2002**, *47*, 3361–3373.
39. C. Bock, A. Smith, B. MacDougall, *Electrochim. Acta* **2002**, *48*, 57–67.
40. G. M. Swain, A. B. Anderson, J. C. Angus, *Mater. Res. Soc. Bull.* **1998**, *23*, 56–60.
41. N. Vinokur, B. Miller, Y. Avyigal et al., *J. Electrochem. Soc.* **1996**, *143*, L238–L240.
42. J. J. Carey, C. S. Christ Jr., S. N. Lowery, Method of electrolysis employing a doped diamond anode to oxidize solutes in wastewater, U.S. Patent 5,399,247, March 21, 1995
43. V. Fischer, D. Gandini, S. Laufer et al., *Electrochimi. Acta* **1998**, *44*, 521–524.
44. G. Foti, D. Gandini, C. Comninellis et al., *Electrochem. Solid-State Lett.* **1999**, *2*, 228–230.
45. A. Perret, W. Haenni, N. Skinner et al., *Diamond Relat. Mater.* **1999**, *8*, 820–823.
46. G. M. Swain, R. Ramesham, *Anal. Chem.* **1993**, *65*, 345–351.
47. N. Weinberg, The Electrosynthesis Co., Inc. [private communication].
48. T. Shimamune, Y. Nakajima, Electrolytic electrode and method of production thereof U.S. Patent 5,683,567, Nov. 4, 1997.
49. A. R. Sacco, T. Shimamune, Y. Nakajima, *Fifth International Forum on Electrolysis in the Chemical Industry*, Electrosynthesis Company, New York, 1991
50. K. Scott, *Bull. Electrochem.* **1993**, *9*, 170–186.
51. G. D. Zappi, S. Arbesman, N. L. Weinberg, *13th International Forum on Electrolysis in the Chemical Industry*, Clearwater Beach, The Electrosynthesis Company, New York, 1999
52. D. F. Steele, *Chem. Br.* **1991**, 915–918.
53. D. F. Steele, D. Richardson, D. R. Craig et al. in *Electrochemistry for a Cleaner*

Environment (Eds.: D. Genders and N. Weinberg), The Electrosynthesis Company, New York, 1992, pp. 287–300; A. C. Almon, B. R. Buchanan, *Electrochemistry for a Cleaner Environment*, The Electrosynthesis Company, New York, 1992, pp. 301–321.
54. J. C. Farmer, F. T. Wang, P. R. Lewis et al., *J. Electrochem. Soc.* **1992**, *139*, 3025–3029.
55. D. Schmal, J. van Erkel, A. M. C. P. de Jong et al. in *Proceedings of 2nd European Conference on Environment Technology*, Amsterdam, The Netherlands (Eds.: K. J. A. de Waal, W. J. Vanden Brink), Martinus Nijhoff Publication, Boston, 1987, pp. 284–293.
56. D. Schmal, P. J. Van Duin, A. M. C. P. Jong, *Dechma-Monographien*, Band 124, VCH Verlagsgesellschalt, 1991, pp. 241–247.
57. D. Schmal, J. Van Erkel, P. J. Van Duin, *J. Appl. Electrochem.* **1986**, *16*, 422–430; *Inst. Chem. Eng. Symp. Ser.* **1986**, *98*, 281–291.
58. P. J. Van Duin, J. Van Erkel, Process for the detoxification of chemical waste materials, US Patent 4,443,309, April 17, 1984
59. B. MacDougall, M. Gattrell, S. Tajwar et al. in *Water Purification by Photocatalytic, Photoelectrochemical and Electrochemical Processes* (Eds.: T. L. Rose, E. Rudd, O. Murphy et al.), The Electrochemical Society, Pennington, 1994, 94–19, pp. 99–110.
60. M. Gattrell, B. MacDougall, *J. Electrochem. Soc.* **1999**, *146*, 3335–3348.
61. B. R. MacDougall, M. A. Gattrell, O. Kargina-Power, Removal of organics from aqueous solutions, US Patent 5,662,789, Sept. 2, 1997
62. T. Guena, L. Wang, M. Gattrell et al., *J. Electrochem. Soc.* **2000**, *147*, 248–255.
63. O. Kargina, B. MacDougall, Yu. M. Kargin et al., *J. Electrochem. Soc.* **1997**, *144*, 3715–3721.
64. E. C. Couture, J. F. Rusling, S. Zhang, *Trans. Inst. Chem. Eng.* **1992**, *70B*, 153–157; see also ICMEME Symposium Series No. 127, **1992**, pp. 177–187.
65. J. F. Rusling, S. Schweizer, S. Zhang et al., *Colloids Surf., A: Physicochem. Eng. Aspects* **1994**, *88*, 41–49.
66. J. F. Rusling, *Microporous Mater.* **1994**, *3*, 1–16.
67. S. Zhang, J. F. Rusling, *Environ. Sci. Technol.* **1995**, *29*, 1195–1199.
68. S. Zhang, J. F. Rusling, *Environ. Sci. Technol.* **1993**, *27*, 1375–1380.
69. J. Schweizer, J. F. Rusling, Q. Huang, *Chemosphere* **1994**, *28*, 961–970.
70. H. Zhang, J. F. Rusling, *Talanta* **1993**, *40*, 741–747.
71. A. Sucheta, I. U. Haque, J. F. Rusling, *Langmuir* **1992**, *8*, 1633–1636.
72. A. E. F. Nassar, J. M. Bobbitt, J. D. Stuart et al., *J. Am. Chem. Soc.* **1995**, *17*, 10986–10993.
73. Z. Zhang, A. E. F. Nassar, Z. Lu et al., *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 1769–1774.
74. K. H. Lee, E. Manlangit, T. C. Franklin et al., *J. Electrochem. Soc.* **1996**, *143*, 3435–3439.
75. T. C. Franklin, J. Darlington, T. Solouski et al., *J. Electrochem. Soc.* **1991**, *138*, 2285–2288.
76. T. C. Franklin, R. Nnodimale, J. Kerimo, *J. Electrochem. Soc.* **1993**, *140*, 2145–2150.
77. R. W. Gillham, S. F. O'Hannesin, *Groundwater* **1994**, *32*, 958–967.
78. R. W. Gillham in *Advances in Groundwater Pollution Control and Remediation* (Ed.: M. M. Aral), Kluwer Academic Publishers, Dordrecht, 1996, 249–274.
79. G. W. Reynolds, T. F. Hoff, R. W. Gillham, *Environ. Sci. Technol.* **1990**, *24*, 135–142.
80. T. L. Johnson, M. M. Scherer, P. G. Tratnyek, *Environ. Sci. Technol.* **1996**, *30*, 2634–2640.
81. R. Focht, J. Vogan, S. F. O'Hannesin, *Remediation* **1996**, *36*, 81–94.
82. United States Environmental Protection Agency, *Field Applications of in-situ Remediation Technologies; Permeable Reactive Barriers*, EPA-542-R-99-002, April 1999
83. J. B. Mosher, L. Figueroa, *Miner. Eng.* **1996**, *9*, 573–582.
84. W. Zabban, R. Helwick, *Plat. Surf. Finish.* **1980**, 56–59.
85. J. Y. Hwang, Y. Y. Wang, C. C. Wan, *J. Appl. Electrochem.* **1987**, *17*, 684–690.
86. J. Feng, D. C. Johnson, *J. Electrochem. Soc.* **1990**, *137*, 507–510.
87. B. Wels, D. C. Johnson, *J. Electrochem. Soc.* **1990**, *137*, 2785–2791.
88. C. S. Hofseth, T. W. Chapman, *J. Electrochem. Soc.* **1999**, *146*, 199–207.
89. F. Hine, M. Yasuda, T. Tida et al., *Electrochim. Acta* **1986**, *31*, 1389–1395.

90. J. J. Byerley, K. Enns, *CIM Bull.* **1984**, *77*, 87–93, No. 861.
91. K. E. Haque, D. J. MacKinnon, *CIM Bull.* **1996**, *89*, 104–106, No. 1001.
92. C. S. Fugivara, P. T. A. Sumodjo, A. A. Castro et al., *The Analyst* **1996**, *121*, 541–546.
93. M. L. Lin, Y. Y. Wang, C. C. Wan, *J. Appl. Electrochem.* **1992**, *22*, 1197–1200.
94. H. Tamura, T. Arikado, H. Yoneyama et al., *Electrochim. Acta* **1974**, *19*, 273–277.
95. T. Arikado, I. Iwakura, H. Yoneyama et al., *Electrochim. Acta* **1976**, *21*, 1021–1027.
96. P. Tissot, M. Fragniere, *J. Appl. Electrochem.* **1994**, *24*, 509–512.
97. D. Pletcher, N. L. Weinberg, *Chem. Eng.* **1992**, 132–141.
98. D. Pletcher, F. C. Walsh in *Electrochemistry for a Cleaner Environment* (Eds.: D. Genders, N. Weinberg), The Electrosynthesis Company, New York, 1992, pp. 51–100.
99. D. Pletcher, F. C. Walsh, *Industrial Electrochemistry*, Chapman & Hall, London and New York, 1990, pp. 336–346.
100. D. Simonsson, *Chem. Soc. Rev.* **1997**, *26*, 181–189.
101. D. Pletcher, F. C. Walsh, *Industrial Electrochemistry*, Chapman & Hall, London and New York, 1990, pp. 346–350.
102. F. C. Walsh, D. R. Gabe, *Trans. Inst. Chem. Eng.* **1990**, *68B*, 107–117.
103. F. C. Walsh, as in ref. 98, pp. 101–159.
104. D. R. Gabe, G. D. Wilson, J. Gonzalez-Garcia et al., *J. Appl. Electrochem.* **1998**, *28*, 759–780.
105. J. M. Marciel, S. M. L. Agostinho, *J. Appl. Electrochem.* **1999**, *29*, 745–750.
106. B. Fleet, *Collect. Czech. Chem. Commun.* **1988**, *53*, 1107–1133.
107. C. J. Brown, as in ref. 98, pp. 183–205.
108. D. Pletcher, as in ref. 98, pp. 11–49.
109. T. A. Davies, as in ref. 98, pp. 173–181.
110. S. Itoi, I. Nakamura, T. Kaahara, *Desalination* **1990**, *32*, 383–389.
111. C. Von Sonntag, G. Mark, R. Martens et al., *J. Water SRT-Aqua* **1993**, *42*, 201–210.
112. K. Rajeshwar, J. Ibanez, *Environmental Electrochemistry*, Academic Press, New York, 1997, Chapter 6.
113. See Chapter 8 in ref. [112].
114. M. I. Litter, *Appl. Catal., B: Environ.* **1999**, *23*, 89–114.
115. A. Heller, J. Schwitzgebel, M. V. Pishko et al. in *Water Purification by Photocatlytic, Photoelectrochemical and Electrochemical Processes* (Eds.: T. L. Rose, O. Murphy, E. Rudd et al.), Electrochemical Society, 1994, 94–19, pp. 1–9.
116. K. Rajeshwar, J. Ibanez, *Environmental Electrochemistry*, Academic Press, New York, 1997, Chapters 6,7 and 8.
117. A. Fujishima, *Abstracts of 1st International Conference On Advanced Oxidation Technologies for Water and Air Remediation*, London, 1994, pp. 98–99.
118. J. Peral, X. Domenechc, D. F. Ollis, *J. Chem. Technol. Biotechnol.* **1997**, *70*, 117–140.
119. A. Heller, *Acc. Chem. Res.* **1995**, *28*, 503–508.
120. N. Serpone, *Res. Chem. Intermed.* **1994**, *20*, 953–992; *Sol. Energy Mater. Sol. C.* **1995**, *38*, 369–379.
121. J. A. Navio, M. C. Hidalgo, G. M. Restrepo, *Spec. Pub. R. Soc. Chem.* **1998**, *217*, 175–188.
122. A. Taghizadeh, M. F. Lawrence, L. Miller et al., *J. Photochem. Photobiol., A: Chem.* **2000**, *130*, 145–156.
123. J. M. Herrmann, *Catal. Today* **1999**, *53*, 115–129.
124. M. Nair, Z. Luo, A. Heller, *Ind. Eng. Chem. Res.* **1993**, *32*, 2318–2323.
125. I. I. Moiseev, *Russ. Chem. Bull.* **1995**, *44*, 560–572.
126. T. A. Davies, J. D. Genders, D. Pletcher, *A First Course in Ion Permeable Membranes*, The Electrochemical Consultancy, Harts, 1997
127. H. Strathmann in *Membrane Handbook* (Eds.: W. S. W. Ho, K. K. Sirkar), Chapman & Hall, New York/London, 1992, pp. 217–262.
128. K. Rajeshwar, J. Ibanez, *Environmental Electrochemistry*, Academic Press, New York, 1997, pp. 432–448; p. 689.
129. H. Strathmann in *Environmental Oriented Electrochemistry* (Ed.: C. A. C. Sequeira), Elsevier, Amsterdam, 1994, pp. 495–533.
130. A. P. Velin, *Sixth International Forum on Electrolysis in the Chemical Industry*, Fort Lauderdale, 1992
131. D. Genders, Watts New-Quarterly Newsletter of the Electrosynthesis Co., Inc., 1995, Vol. 1, 1; website at www.electrosynthesis.com.

132. D. Pletcher, J. D. Genders, N. L. Weinberg et al., Electrochemical methods for production of alkali metal hydroxides without the co-production of chlorine, U.S. Patent 5,246,551, Sept. 21, 1993.
133. K. Mani, *J. Membrane Sci.* **1991**, *58*, 117–138.
134. K. Nagasubramanian, F. P. Chlanda, K. J. Liu, *J. Membrane Sci.* **1977**, *2*, 109–124.
135. B. Bauer, F. J. Gerner, H. Strathmann, *Desalination* **1988**, *68*, 279–292.
136. C. P. Jones, M. D. Neville, A. D. Turner in *Electrochemistry for a Cleaner Environment* (Eds.: J. D. Genders, N. Weinberg), The Electrosynthesis Company., East Amherst, 1992, pp. 207–220.
137. P. M. Allen, P. G. Griffiths, C. P. Jones et al., *Spec. Pub. R. Soc. Chem.* **1996**, *182*, 96–103.
138. I. W. Cumming, H. Tai, M. Beier, *Chem. Eng. Res. Des.* **1997**, *75*, 9–13.
139. F. DiMascio, J. Wood, J. M. Fenton, *Interface*, The Electrochemical Society, Pennington, 1998, pp. 26–31, Vol. 7, No. 3.
140. G. C. Ganzi, A. D. Jna, F. DiMascio et al., *Ultrapure Water* **1997**, *14*, 64–69.
141. H. Neumeister, R. Flucht, L. Furst et al., *Ultrapure Water* **2000**, *17*, 22–30.
142. Y. B. Acar, A. N. Alshawabkeh, *Environ. Sci. Technol.* **1993**, *27*, 2638–2647.
143. R. J. Gale, H. Li, Y. B. Acar in *Environmentally Oriented Electrochemistry* (Ed.: C. A. C. Sequeira), Elsevier, Amsterdam, 1994, pp. 621–654.
144. R. F. Probstein, R. E. Hicks, *Science* **1993**, *260*, 498–504.
145. R. Clarke, *Ninth International Forum on Electrolysis in the Chemical Industry*, Clearwater Beach, Electrosynthesis Company, New York, 1995
146. R. Lageman, P. Wieberen, G. Seffinga, *Chem. Ind. (London)* **1989**, *9*, 585–590.
147. Y. B. Acar, R. J. Gale, A. N. Alshawabkeh et al., *J. Hazard. Mater.* **1995**, *40*, 117–138.
148. K. Rajeshwar, J. Ibanez, *Environmental Electrochemistry*, Academic Press, New York, 1997, p. 452.
149. S. V. Ho, C. J. Athmer, P. W. Sheridan et al., *J. Hazard. Mater.* **1997**, *55*, 39–60.
150. S. V. Ho, C. Athmer, P. W. Sheridan et al., *Environ. Sci. Technol.* **1999**, *33*, 1086–1099.
151. D. T. Mah, J. W. Weidner, S. Motupally, *J. Electrochem. Soc.* **1999**, *146*, 3924–3947.
152. P. Ribordy, C. Pulgarin, J. Kiwi et al., *Water. Sci. Technol.* **1997**, *35*, 293–302.
153. J. Janata, *Anal. Chem.* **1992**, *64*, 196R-.
154. R. I. Stefan, J. F. van Staden, H. Y. Aboul-Enein, *Crit. Rev. Anal. Chem.* **1999**, *29*, 133–153.
155. J. P. Hart, S. A. Wring, *Trends Anal. Chem.* **1997**, *16*, 89–103.
156. S. Saini, A. P. F. Turner, *Trends Anal. Chem.* **1995**, *14*, 304–310.
157. B. J. Seddon, Y. Shao, H. H. Girault, *Electrochim. Acta* **1994**, *39*, 2377–2386.
158. J. Janata, M. Josowicz, P. Vanysek et al., *Anal. Chem.* **1998**, *70*, R179–R208.
159. See Chapter 4 in Ref. [6].